# THE ANNALS OF MATHEMATICAL STATISTICS

Editor
T. W. ANDERSON

Associate Editors
M. A. GIRSHICK
HERMAN RUBIN
ALEXANDER M. MOOD
JOHN W. TUKEY

Editorial Board
M. S. BARTLETT
J. W. BRUNSCHWIG
DAVID BLACKWELL
HARALD CRAMÉR
WILLIAM G. COCHRAN
J. L. DOOB
W. E. DEMING
[illegible]
[illegible]
[illegible]

# THE $\chi^2$ TEST OF GOODNESS OF FIT[1]

By William G. Cochran

*Johns Hopkins University*

**1. Summary.** This paper contains an expository discussion of the chi square test of goodness of fit, intended for the student and user of statistical theory rather than for the expert. Part I describes the historical development of the distribution theory on which the test rests. Research bearing on the practical application of the test—in particular on the minimum expected number per class and the construction of classes—is discussed in Part II. Some varied opinions about the extent to which the test actually is useful to the scientist are presented in Part III. Part IV outlines a number of tests that have been proposed as substitutes for the chi square test (the $\omega^2$ test, the smooth test, the likelihood ratio test) and Part V a number of supplementary tests (the run test, tests based on low moments, subdivision of chi square into components).

**2. Introduction.** In the standard applications of the test, the $n$ observations in a random sample from a population are classified into $k$ mutually exclusive classes. There is some theory or null hypothesis which gives the probability $p_i$ that an observation falls into the $i$th class ($i = 1, 2, \cdots, k$). Sometimes the $p_i$ are completely specified by the theory as known numbers, and sometimes they are less completely specified as known functions of one or more parameters $\alpha_1, \alpha_2, \cdots$ whose actual values are unknown. The quantities $m_i = np_i$ are called the *expected* numbers, where

$$\sum_{i=1}^{k} p_i = 1, \qquad \sum_{i=1}^{k} m_i = n.$$

The starting point in the theory is the joint frequency distribution of the *observed* numbers $x_i$ falling in the respective classes. If the theory is correct, these observed numbers follow a multinomial distribution with the $p_i$ as probabilities. The joint distribution of the $x_i$ is therefore specified by the probabilities

$$\frac{n!}{x_1!\, x_2! \cdots x_k!}\, p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}. \tag{1}$$

As a test criterion for the null hypothesis that the theory is correct, Karl Pearson [1] proposed the quantity

$$X^2 = \sum_{i=1}^{k} \frac{(x_i - m_i)^2}{m_i} = \sum_{i=1}^{k} \frac{x_i^2}{m_i} - n. \tag{2}$$

---

[1] Department of Biostatistics Paper No. 282.

*Editor's Note:* This paper was presented to the Boston meeting of the Institute of Mathematical Statistics, December 28, 1951, and is published in the *Annals* by invitation of the Institute Committee on Special Invited Papers.

Pearson did not mention any particular alternative hypothesis. The test has usually been regarded as applicable to situations in which the alternative hypothesis is described only in rather vague and general terms.

As with any test of a hypothesis, certain properties of the test must be worked out before it is ready for practical application. We need to know the frequency distribution of the test criterion when the null hypothesis is correct, in order that tables of significant values can be constructed. As much as possible should also be known about the performance of the test when the null hypothesis does not hold. Practically all the results in the literature deal only with the limiting distribution of $X^2$ as $n \to \infty$, the $p_i$ remaining fixed. When the null hypothesis holds, this limiting distribution is the $\chi^2$ distribution,

$$(3) \qquad \frac{1}{2^{\nu/2}\left(\frac{\nu}{2}-1\right)!}\,(\chi^2)^{(\nu/2)-1}e^{-\frac{1}{2}\chi^2}\,d\chi^2,$$

where $\nu$ is the number of degrees of freedom in $\chi^2$. This distribution is also known to be that followed by the quantity

$$y_1^2 + y_2^2 + \cdots + y_\nu^2,$$

where the $y_i$ are normally and independently distributed with zero means and unit variances.

To avoid confusion, the symbol $X^2$ will be used for the quantity in equation (2) which is calculated from the data when a chi square test is performed. The symbol $\chi^2$ will refer to any random variate which follows the tabular chi square distribution given in (3).

## Part I. Historical Development of the Test

**3. Karl Pearson's 1900 paper.** This remarkable paper is one of the foundations of modern statistics. Its style has always impressed me as unusual for a pioneering paper. Pearson writes with the air of a man who knows exactly what he is doing. The exposition, although clear, is slightly hurried and brusque, as if the reader will not wish to be troubled by elaborate details of a problem that is routine and straightforward. One misses any discussion of how Pearson came to choose the $X^2$ test criterion, and of when he first came to realize that this criterion would, under certain circumstances, follow the $\chi^2$ distribution.

The paper opens by proving that if a set of $\nu$ correlated variates $z_i$, with zero means, follow a multivariate normal distribution

$$Ce^{-\frac{1}{2}Q}dz_1dz_2\cdots dz_\nu,$$

then the quadratic form $Q$ is distributed as $\chi^2$ with $\nu$ degrees of freedom. This proof is accomplished, in about half a page, by a now familiar geometrical argument. Pearson points out that the ellipsoid $Q$ can be "squeezed" into a sphere. A transformation to polar coordinates is made, where $\chi$ is the radius of the sphere and $Q = \chi^2$. He then remarks that all the angles introduced in

the transformation will integrate out to a constant factor, so that the probability that $Q$ exceeds $\chi_0^2$, say, reduces to

$$P = \frac{\int_{\chi_0}^{\infty} e^{-\frac{1}{2}\chi^2}\chi^{\nu-1}\,d\chi}{\int_0^{\infty} e^{-\frac{1}{2}\chi^2}\chi^{\nu-1}\,d\chi}.$$

This, of course, is the tabular $\chi^2$ distribution, expressed as an integral of $\chi$ rather than $\chi^2$. The result is a generalization of the result reached by Helmert in 1876, and also of the result which Student later developed in 1908 as groundwork for the $t$-distribution.

The next step is to express the probability integral in power series form, this being necessary to construct a table of the probability integral. The paper contains a table giving $P$ to 6 decimal places, for degrees of freedom from 2 to 19, and for various integral values of $\chi^2$.

Pearson now turns to the problem of testing goodness of fit. He deals first with the case in which the expectations $m_i$ are known numbers. The data have been classified as described in Section 2, so that the observations $x_i$ follow a multinomial distribution. Pearson assumes without more ado that the $x_i$ may be taken as normally distributed. It is at this point, therefore, that he is committed to the assumption that the expectations $m_i$ are large in all cells. He assigns to the $x_i$ their correct variances and covariances from the multinomial distribution, that is,

$$\sigma_{ii} = np_i(1 - p_i), \qquad \sigma_{ij} = -np_ip_j \qquad (i \neq j) \tag{4}$$

The remainder of the proof consists in writing down the presumed multivariate normal distribution of the quantities $(x_i - m_i)$. From this comes the pleasing result that the quadratic form $Q$ in the exponent is simply

$$Q = \sum \frac{(x_i - m_i)^2}{m_i} = X^2.$$

This may be shown as follows. Since the $x_i$ are constrained to add to $n$, we must omit one of them, say $x_k$, in considering their joint distribution. If the joint frequency function of the first $(k - 1)$ of the $x$'s is $Ce^{-\frac{1}{2}Q}$, it is well known that

$$Q = \sum_{i=1}^{k-1}\sum_{j=1}^{k-1} \sigma^{ij}(x_i - m_i)(x_j - m_j),$$

where $\sigma^{ij}$ is the inverse of the matrix $\sigma_{ij}$ given in (4).

Now consider $X^2$, with $(x_k - m_k)$ replaced by $-\sum_{i=1}^{k-1}(x_i - m_i)$.

$$X^2 = \frac{(x_1 - m_1)^2}{m_1} + \cdots + \frac{(x_{k-1} - m_{k-1})^2}{m_{k-1}} + \frac{\{(x_1 - m_1) + \cdots + (x_{k-1} - m_{k-1})\}^2}{m_k}.$$

Hence, if we write

$$X^2 = \sum_{i=1}^{k-1}\sum_{j=1}^{k-1} a_{ij}(x_i - m_i)(x_j - m_j),$$

the matrix $a_{ij}$ is

$$a_{ii} = \frac{1}{m_i} + \frac{1}{m_k}, \qquad a_{ij} = \frac{1}{m_k} \qquad (i \neq j). \tag{5}$$

The remainder of the proof consists in showing that (5) is the inverse of (4). Pearson does this by a rather complicated polar transformation, but the student who has some familiarity with the evaluation of determinants will find it a fairly easy exercise, as Hotelling [2] has pointed out. It may be helpful to write (4) as

$$\sigma_{ii} = m_i\left(1 - \frac{m_i}{n}\right), \qquad \sigma_{ij} = -\frac{m_i m_j}{n} \qquad (i \neq j), \tag{4'}$$

and to invert (5) rather than (4), or to prove that the product of the matrices $a_{ij}$, $\sigma_{jk}$ is the unit matrix.

Hence, by the first part of Pearson's paper, we reach the result that in the limit as $n$ becomes large, $X^2$ follows the $\chi^2$ distribution with $(k-1)$ degrees of freedom.

An approach which avoids most of the mathematical complexities in Pearson's argument has been pointed out by Fisher [3]. If the observations $x_i$ are regarded as following independent Poisson distributions, their joint frequency function is

$$\prod_{i=1}^{k} \frac{e^{-m_i} m_i^{x_i}}{x_i!} = e^{-n} \prod_{i=1}^{k} \frac{m_i^{x_i}}{x_i!}, \tag{6}$$

since $\sum m_i = n$.

Under this assumption, their total $T = \sum x_i$ also follows a Poisson distribution, with mean $\sum m_i = n$. The frequency function of $T$ is therefore

$$\frac{e^{-n} n^T}{T!}. \tag{6'}$$

Hence, on dividing (6) by (6') the *conditional* frequency function of the $x_i$, given that their total $T$ has the value $n$, is

$$\frac{n!}{x_1!\, x_2! \cdots x_k!}\left(\frac{m_1}{n}\right)^{x_1}\left(\frac{m_2}{n}\right)^{x_2} \cdots \left(\frac{m_k}{n}\right)^{x_k}.$$

This is the same as the basic multinomial (1).

This argument implies that in an investigation of the distribution of $X^2$ we may start by regarding the $x_i$ as following independent Poisson distributions, subject to the restriction that $\sum x_i = n$.

In the limit, as the $m_i$ become large, the quantities

$$y_i = \frac{x_i - m_i}{\sqrt{m_i}}$$

become normally distributed with means zero and unit standard deviations, since the Poisson distribution of $x_i$ has mean $m_i$ and standard deviation $\sqrt{m_i}$. Hence the limiting distribution of $X^2$ is that of the quantity

$$y_1^2 + y_2^2 + \cdots + y_k^2,$$

where the $y_i$ are independently distributed but are subject to the single linear restriction

$$\sum_{i=1}^{k} y_i \sqrt{m_i} = \sum_{i=1}^{k} (x_i - m_i) = 0.$$

The fact that in the limit $X^2$ follows the $\chi^2$ distribution with $(k - 1)$ degrees of freedom can now be established by integration or by quoting well known theorems on the analysis of variance. This approach also makes it clear that if further homogeneous linear restrictions are imposed on the variates $(x_i - m_i)$, either by the structure of the data or in the process of fitting, the effect will merely be to reduce the degrees of freedom in $\chi^2$.

Pearson next considers the situation in which the $m_i$ depend on parameters that have to be estimated from the sample. Denoting by $m_i'$ the expectations derived from sample estimates of these parameters, and by $m_i$ the true expectations, he discusses the difference

$$X^2 - X'^2 = \sum_{i=1}^{k} \frac{x_i^2}{m_i} - \sum_{i=1}^{k} \frac{x_i^2}{m_i'}.$$

He suggests that this difference will usually be positive, because we ought to be able to do a better job of fitting when we can adjust the estimates of the parameters to suit the vagaries of the sample. He argues, however, that the difference will be small enough so that if we regard $X'^2$ as also distributed as $\chi^2$ with $(k - 1)$ degrees of freedom, the error in this approximation will not affect practical decisions.

In this conclusion, which is reached with some sign of hesitation, he may well have been justified for many applications. We now know that the number of degrees of freedom must be reduced in order to give the correct limiting distribution. Perhaps the most common of all uses of the $X^2$ test is for the $2 \times 2$ contingency table. Unfortunately, Pearson's suggestion works rather poorly in this case, since he attributed 3 degrees of freedom to $X^2$, whereas it should receive only 1. This point caused some confusion and controversy in practical applications, and was not settled for over 20 years.

Finally, the paper contains eight numerical applications of the new technique. In two of these he pokes fun at Sir George Airy and Professor Merriman. They had both published series of observations which they claimed to be good illus-

trations of variates that follow the normal distribution. In the absence of any test of goodness of fit, Airy and Merriman could judge this question only by eye inspection. Pearson showed that the significance probability for Airy's data was 0.014, although the data from which Pearson calculated $X^2$ had already been smoothed by Airy. Merriman fared worse, his probability being $1\frac{1}{2}$ parts in a million. These examples show the weak position in which the scientist was placed when he had to judge goodness or badness of fit in the absence of an objective test of significance.

To summarize, Pearson established the necessary distribution theory for finding significance levels when the null hypothesis provides the exact values of the $m_i$, except that he did not show that the exact distribution of $X^2$, which is discontinuous, actually approaches $\chi^2$ as a limiting distribution. A fully rigorous proof may be given by the use of moment-generating functions [8].

**4. The distribution of $X^2$ when the expectations are estimated from the sample.** This problem is much more difficult and was not elucidated until the appearance of Fisher's 1924 paper. In the intervening period, a paper by Greenwood and Yule [4] in 1915 illustrates the perplexity which existed among critical users of the test and which led to the "degrees of freedom" battle. The authors were attempting to examine the effects of inoculation against typhoid and cholera. They present a substantial number of $2 \times 2$ tables containing subjects classified as inoculated or not, and also as to whether they contracted the disease following exposure to it. The following is an example.

*Kalain (Cholera)*

| | Not Attacked | Attacked | Total |
|---|---|---|---|
| Inoculated | 1625 | 5 | 1630 |
| Not | 1022 | 11 | 1033 |
| Total | 2647 | 16 | 2663 |

$X^2 = 6.08$.

Following Pearson's rule, they assign 3 degrees of freedom to $X^2$. This gives a $P$ of 0.108, whereas with 1 degree of freedom, $P$ is 0.015. They realised, however, that the hypothesis that inoculation is without effect could be tested, alternatively, by calculating the difference $(p_1 - p_2)$ between the percent ill among the inoculated and the non-inoculated. On the null hypothesis, the ratio

$$R = \frac{p_1 - p_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

is approximately a normal deviate with mean zero and unit standard deviation. This test, as they found, gave more statistically significant results than Pear-

son's test. The quantity $R$ is exactly equal to $X$ if we use the pooled percent ill, $\bar{p}$, in estimating the two variances in the denominator, so that the "normal deviate" test and the $X^2$ test should be identical. It is not clear that Greenwood and Yule recognised this in 1915.

Although giving the impression of being somewhat in a quandary as to which test to employ, they content themselves with the decision to adopt Pearson's test, pointing out that it is the more conservative of the two, and adding that the issue deserves further theoretical investigation.

After some controversy, the matter was cleared up in theoretical papers [3], [5] by Fisher in 1922 and 1924, supported by sampling experiments which were published by Yule [6] and Brownlee [7]. Fisher's 1922 paper included a discussion of $2 \times 2$ contingency tables, and showed that for this case $X^2$ is the square of a single quantity which had a limiting normal distribution, and that the $X^2$ test and the test of $(p_1 - p_2)$ by the normal distribution are identical.

Fisher's 1924 paper is much more general. He points out that the limiting distribution of $X^2$ depends on the method of estimation. With a poor method of estimation, $X^2$ may frequently have a large value even if the theory is correct. It is therefore necessary, in a general proof of the distribution of $X^2$, to state what is to be the method of estimation. At first sight, the natural method would seem to be to choose the unknown parameters so that $X^2$ is as small as possible. Fisher shows that in the limit in large samples, this method becomes equivalent to the method of maximum likelihood. For his main proof, this result serves as an ingenious lemma, since at one point in the main proof he assumes that estimation is by maximum likelihood, while at another he assumes that it is by minimum $X^2$.

Although Fisher's main proof is not fully rigorous, it is worthwhile to outline the principal steps, because the proof does reveal the core of the problem, and a rigorous proof requires advanced methods. Fisher starts in the same way as Pearson, by considering

$$X^2 - X'^2 = \sum_{i=1}^{k} \frac{x_i^2}{m_i} - \sum_{i=1}^{k} \frac{x_i^2}{m_i'} = \sum_{i=1}^{k} x_i^2 \left( \frac{1}{m_i} - \frac{1}{m_i'} \right),$$

where $m_i$ is a specified function of a single unknown parameter $\alpha$, with $m_i = m_i(\alpha)$, $m_i' = m_i(\alpha')$. He expands in a Taylor series about the point $\alpha'$. The first two terms give

$$X^2 - X'^2 = -\sum \frac{x_i^2}{m_i'^2} \left( \frac{\partial m_i'}{\partial \alpha'} \right) \delta\alpha + \sum x_i^2 \left\{ \frac{2}{m_i'^3} \left( \frac{\partial m_i'}{\partial \alpha'} \right)^2 - \frac{1}{m_i'^2} \frac{\partial^2 m_i'}{\partial \alpha'^2} \right\} \frac{(\delta\alpha)^2}{2!},$$

where $\delta\alpha = (\alpha - \alpha')$. Since the method of estimation consists in choosing $\alpha'$ so that $X'^2$ is a minimum, we have

$$\sum \frac{x_i^2}{m_i'} \left( \frac{\partial m_i'}{\partial \alpha'} \right) = 0,$$

so that the first term on the right vanishes.

In the second term on the right, Fisher replaces $x_i$ by $m_i'$. The error intro-

duced by this step may be shown to be of the same order as the third term in the Taylor series, which has already been ignored. Hence,

$$X^2 - X'^2 = \sum \left\{ \frac{2}{m_i'} \left( \frac{\partial m_i'}{\partial \alpha'} \right)^2 - \frac{\partial^2 m_i'}{\partial \alpha'^2} \right\} \frac{(\delta\alpha)^2}{2!}.$$

But if the identity $\sum m_i' = n$ is differentiated twice, we find

$$\sum \left( \frac{\partial^2 m_i'}{\partial \alpha'^2} \right) = 0.$$

Hence,

$$X^2 - X'^2 = \sum \left\{ \frac{1}{m_i'} \left( \frac{\partial m_i'}{\partial \alpha'} \right)^2 \right\} (\delta\alpha)^2.$$

If $\alpha'$ is regarded as a maximum likelihood estimate of $\alpha$, we may use the standard result that the error of estimate $(\alpha' - \alpha)$ has a limiting normal distribution, with mean zero, and variance given by

$$\frac{1}{\sigma^2_{\alpha'}} = \sum \left\{ \frac{1}{m_i} \left( \frac{\partial m_i}{\partial \alpha} \right)^2 \right\} \doteq \sum \left\{ \frac{1}{m_i'} \left( \frac{\partial m_i'}{\partial \alpha'} \right)^2 \right\}.$$

This gives the neat result

$$X^2 - X'^2 = \frac{(\alpha' - \alpha)^2}{\sigma^2_{\alpha'}}.$$

Our object is to find the limiting distribution of $X'^2$. At this point the facts in our possession are: (i) $X^2$ is distributed as $\chi^2$ with $(k - 1)$ degrees of freedom (this follows from Pearson's results, since $X^2$ is calculated from the correct $m$'s); and (ii) $X^2 - X'^2$ is distributed as $\chi^2$ with 1 degree of freedom (from Fisher's argument). These facts are not sufficient to determine the distribution of $X'^2$. However, Fisher points out that the limiting distributions of $X'^2$ and $(\alpha' - \alpha)^2$ must be independent, since $X'^2$ was obtained by minimizing $X^2$ with respect to $\alpha'$. Given this additional result, it is easily shown that $X'^2$ must be distributed as $\chi^2$ with $(k - 2)$ degrees of freedom.

The argument leads to two further results. Any method of estimation that is efficient gives estimates which become, in the limit, identical with the maximum likelihood estimate. Thus the $\chi^2$ distribution, with the appropriate reduction in degrees of freedom, is valid for any efficient method of estimation. The argument also provides the limiting mean value of $X'^2$ when the estimation is inefficient. An interesting corollary is that the mean value of $X'^2$ exceeds that of $X^2$ when the efficiency is less than 50 percent.

Rigorous proofs of the general limiting distribution, when several parameters are being estimated, are scarce in the literature. For the student, one of the best is that given by Cramér [8]. The restrictions under which he proves his result are stated below. He assumes maximum likelihood estimation.

THEOREM. *Suppose that the $k$ probabilities $p_i(\alpha_1, \alpha_2, \cdots, \alpha_s)$ are known functions of $s < k$ parameters $\alpha_1, \alpha_2, \cdots, \alpha_s$. For all points of a nondegenerate interval $A$ in the $s$-dimensional space of the $\alpha_j$, the $p_i$ satisfy the following conditions*:

(a) $$\sum_{i=1}^{k} p_i(\alpha_1, \cdots, \alpha_s) = 1;$$

(b) $$p_i(\alpha_1, \cdots, \alpha_s) > C^2 > 0 \qquad \text{for all } i;$$

(c) *every $p_i$ has continuous derivatives* $\dfrac{\partial p_i}{\partial \alpha_j}$ *and* $\dfrac{\partial^2 p_i}{\partial \alpha_j \partial \alpha_h}$;

(d) *the matrix* $D = \left(\dfrac{\partial p_i}{\partial \alpha_j}\right)$ *is of rank $s$. Then $X^2$ is distributed in the limit, as $n \to \infty$, in a $\chi^2$ distribution with $(k - s - 1)$ degrees of freedom.*

**5. The limiting power function of the test.** The literature does not contain much discussion of the power function of the $X^2$ test. There has been little demand for this from applications, because the test is most commonly used when we do not have a clear-cut alternative in mind, and are not in a position to make computations of the power.

Suppose that we test the null hypothesis that the expectations are $m_i$ when in fact they are $m_i'$. If the values of $m_i$, $m_i'$ and the significance level are kept fixed, then as $n$ increases, it turns out, as would be expected, that the power of the test tends to 1. This has been shown by Neyman [9]. In order to examine the situation in which the power is not close to 1 in large samples, we must somehow make the task continually harder for the test as $n$ becomes larger. This can be accomplished either by making the significance probability decrease steadily as $n$ increases, thus reducing the chance of an error of type I, or by moving the alternative hypothesis steadily closer to the null hypothesis. The second method will be discussed here. Let

$$m_i' - m_i = c_i\sqrt{n}; \qquad \text{that is, } p_i' - p_i = c_i/\sqrt{n},$$

where the quantities $c_i$ remain fixed as $n$ increases.

The nature of the limiting power distribution of $X^2$ is indicated by the following argument, for which I am indebted to J. W. Tukey. We may write

$$\frac{x_i - m_i}{\sqrt{m_i}} = \frac{(x_i - m_i')}{\sqrt{m_i'}}\sqrt{\frac{m_i'}{m_i}} + \frac{m_i' - m_i}{\sqrt{m_i}}. \tag{7}$$

Now

$$\sqrt{\frac{m_i'}{m_i}} = \sqrt{1 + \frac{m_i' - m_i}{m_i}} = \sqrt{1 + \frac{c_i}{p_i\sqrt{n}}}.$$

This tends to 1 as $n$ becomes large, since $c_i$ and $p_i$ are presumed to remain fixed. If we adopt Fisher's approach to the distribution theory (Section 3), the quantities

$$\frac{x_i - m_i'}{\sqrt{m_i'}}$$

tend to become normally and independently distributed with means zero and unit standard deviations, as $n$ becomes large. Consequently, so do the quantities

$$y_i = \frac{(x_i - m_i')}{\sqrt{m_i'}} \sqrt{\frac{m_i'}{m_i}}.$$

Finally, by equation (7),

$$X^2 = \sum_{i=1}^{k} \frac{(x_i - m_i)^2}{m_i} = \sum_{i=1}^{k} \left\{ y_i + \frac{m_i' - m_i}{\sqrt{m_i}} \right\}^2 = \sum_{i=1}^{k} (y_i + a_i)^2,$$

where the $y_i$ are subject to the linear restriction

$$\sum_{i=1}^{k} y_i \sqrt{m_i} = \sum_{i=1}^{k} (x_i - m_i') = 0.$$

Thus, in the limit, $X^2$ is distributed as a sum of squares of variates $(y_i + a_i)$ independently and normally distributed with unit variances, but where the means $a_i$ are not all zero. The variates are subject to one linear restriction when the $m_i$ are completely specified.

This type of distribution is known as a noncentral $\chi^2$. It depends on two parameters—the degrees of freedom, in this case $(k - 1)$, and a parameter of non-centrality $(a_1^2 + a_2^2 + \cdots + a_k^2)$, which has the value

$$\sum_{i=1}^{k} \frac{(m_i' - m_i)^2}{m_i} = \sum_{i=1}^{k} \frac{c_i^2 n}{p_i n} = \sum_{i=1}^{k} \frac{c_i^2}{p_i}.$$

Tables of the noncentral distribution have been provided by Fix [10] and approximations studied by Patnaik [11].

When the $m_i$ have to be estimated from the data, the limiting noncentral $\chi^2$ distribution still holds, with a reduced number of degrees of freedom. A rigorous proof is obtained from Wald's derivation of the limiting distribution of the likelihood ratio test criterion [12], which becomes equivalent to $X^2$ in large samples.

**6. Conditional $X^2$ tests.** As has been mentioned, additional homogeneous linear restrictions imposed on the deviations $(x_i - m_i)$ have the effect of reducing the number of degrees of freedom attributed to $\chi^2$ in the limiting distribution of $X^2$. These restrictions may arise in the process of fitting, or by the nature of the data. They may also be deliberately imposed by the statistician in the device known as a conditional test. This device is illustrated by the $2 \times 2$ contingency table, in which it has created some stimulating discussion [13].

The data are classified according to two different axes, $A$ and $B$.

| | $B_1$ | $B_2$ | Totals |
|---|---|---|---|
| $A_1$ | $x_{11}$ | $x_{12}$ | $r_1$ |
| $A_2$ | $x_{21}$ | $x_{22}$ | $r_2$ |
| Totals..... | $c_1$ | $c_2$ | $n$ |

Data of this kind occur in at least three distinct experimental situations.

(*i*) We select a random sample of $n$ from some population and classify every observation into one of the four cells. The symbol $x_{ij}$ denotes the observed number falling in class $A_iB_j$, while $p_{ij}$ denotes the corresponding probability of falling in this class, where the sum of the four $p$'s is unity. The null hypothesis that the two classifications are *independent* amounts to the relation

$$p_{11}/p_{12} = p_{21}/p_{22}. \tag{8}$$

The joint probability of this group of observations is the usual multinomial

$$\frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}\,p_{11}^{x_{11}}\,p_{12}^{x_{12}}\,p_{21}^{x_{21}}\,p_{22}^{x_{22}}. \tag{9}$$

Only two of the $p_{ij}$ need to be estimated from the data, because of equation (8) and the fact that the $p_{ij}$ add to 1. Thus $X^2$ has $(4 - 2 - 1)$ or 1 degree of freedom.

(*ii*) We take a random sample of size $r_1$ from a population denoted by $A_1$, and an *independent* random sample of size $r_2$ from another population denoted by $A_2$. The null hypothesis states that the probability $p$ of an observation falling in $B_1$ is the same in both populations $A_1$ and $A_2$. Given the null hypothesis, the probability of the sample is the product of the two binomials

$$\left\{\frac{r_1!}{x_{11}!\,x_{12}!}\,p^{x_{11}}q^{x_{12}}\right\}\left\{\frac{r_2!}{x_{21}!\,x_{22}!}\,p^{x_{21}}q^{x_{22}}\right\}. \tag{10}$$

This is not the same as the multinomial (9). Given data of type (*i*), however, let us arbitrarily impose the restriction that in repeated sampling we will consider only those tables which have the same marginal totals $r_1$, $r_2$ as our data. Then (9) must be replaced by the conditional distribution of the $x_{ij}$, given $r_1$ and $r_2$. This conditional distribution is easily seen to be the same as (10). For, starting with (9), the distribution of $r_1$ (and hence $r_2$) is the binomial

$$\frac{n!}{r_1!\,r_2!}\,(p_{11} + p_{12})^{r_1}(p_{21} + p_{22})^{r_2}. \tag{11}$$

To obtain the conditional distribution, we divide (9) by (11). The quotient reduces to (10) if we note that from (8),

$$\frac{p_{11}}{p_{11} + p_{12}} = \frac{p_{21}}{p_{21} + p_{22}} = p \quad \text{(say)}.$$

(*iii*) A third case is obtained if *both* sets of marginal totals are regarded as fixed in repeated sampling. Fisher's tea-tasting experiment [14] is an example. The $A$ classification tells whether the milk or the tea was added first, and the $B$ classification tells whether the lady guessed that the milk or the tea was added first. In Fisher's original experiment, he recommended that the lady be informed how many cups were of each kind, and pointed out that she would presumably match her guesses to those two numbers. Thus in repeated trials it is natural to regard both sets of margins as fixed.

The appropriate basic distribution of the $x_{ij}$ is the conditional distribution which develops from (10) if we keep $c_1$ and $c_2$ fixed. This is found to be

$$\frac{r_1!\, r_2!\, c_1!\, c_2!}{n!\, x_{11}!\, x_{12}!\, x_{21}!\, x_{22}!}. \tag{12}$$

Case (*i*) has 2 unknown parameters and 1 linear restriction on the $x_{ij}$ ; case (*ii*) has 1 unknown parameter and 2 restrictions, while case (*iii*) has no unknown parameters and 3 restrictions.

Is the same $X^2$ test to be used for all cases? In large samples there is no conflict, because $X^2$ has the same limiting distribution however the linear restrictions arise. This is not so in small samples, where the distribution of $X^2$ differs in the three cases. Fisher [15] recommends that the distribution of $X^2$ obtained in case (*iii*) be taken as the exact small-sample distribution for all three types of data. Questions have been raised about this recommendation.

Originally, part of the objection came perhaps from a feeling that there is something improper in keeping the marginal totals fixed in cases (*i*) and (*ii*), because if we actually drew repeated samples of the same size by the same methods, the margins would not all remain fixed. For a rational appraisal, however, the only relevant factors are the effects of the marginal restrictions on the significance probabilities (or type I errors) and on the power (or type II errors). As regards type I errors, Fisher's recommendation has the great advantage that in case (*iii*) the significance probabilities can be computed exactly, whereas in cases (*i*) and (*ii*) the distribution of $X^2$ involves nuisance parameters, so that "exact" probabilities are not available.

The issue thus reduces to the question whether any loss of power occurs if the case (*iii*) test is employed with the first two types of data. For case (*ii*) data, Barnard [13] proposed a different test which in some circumstances appeared to give a small increase in power. More recently K. D. Tocher [16], has proved the remarkable result that a modification of Fisher's test is the most powerful, in the sense of Neyman and Pearson, for one-tailed tests with any

of the three types of data. The modification is necessary to make the problem amenable to Neyman and Pearson's techniques.

The modification may be illustrated by the example which Tocher presents.

TABLE 1

*Tocher's illustration*

Original table

| | | |
|---|---|---|
| 2 | 5 | 7 |
| 3 | 2 | 5 |
| 5 | 7 | 12 |

More extreme cases

| | | |
|---|---|---|
| 1 | 6 | 7 |
| 4 | 1 | 5 |
| 5 | 7 | 12 |

| | | |
|---|---|---|
| 0 | 7 | 7 |
| 5 | 0 | 5 |
| 5 | 7 | 12 |

Given the data on the left, we wish to make a one-tailed test at the 5% level. The two possible sets of data which deviate more from the null hypothesis are shown on the right. In Fisher's exact test, we add the probabilities of the three tables as computed by formula (12). This gives

$$0.26515 + 0.04399 + 0.00126 = 0.31040.$$

This value is regarded as the significance probability.

In Tocher's modification, we also compute the total probability of all more extreme cases, that is,

$$0.04399 + 0.00126 = 0.04525.$$

If these numbers, 0.31040 and 0.04525, are both *below* the stated significance level, 0.05, we reject the hypothesis. If they are both *above* 0.05, we accept. If one is above and one is below, as in the present example, we calculate the ratio

$$\frac{0.05 - 0.04525}{0.26515} = 0.01791.$$

Now draw a random number between 0 and 1. If this number is less than 0.01791, we reject; if greater, we accept.

Although this procedure may appear somewhat startling at first sight, the idea is basically simple. Consider how we can obtain a one-tailed test at the 0.05 level from the $2 \times 2$ table in this example. If the null hypothesis is rejected only when the two most extreme cases on the right of Table 1 occur, the significance level is actually 0.04525. The third most extreme case, represented by the data on the left of Table 1, occurs with probability 0.26515. Consequently, by the computation above, we obtain a test at the 0.05 level if we also declare as significant a fraction 0.01791 of the cases in which the data on the left are encountered. Tocher selects this fraction by a table of random numbers. There seems no other logical basis for deciding which particular fraction to select.

## Part II. Some Aspects of the Practical Use of the Test

**7. The minimum expectation.** Since $\chi^2$ has been established as the limiting distribution of $X^2$ in large samples, it is customary to recommend, in applications of the test, that the smallest expected number in any class should be 10 or (with some writers) 5. If this requirement is not met in the original classification, combination of neighboring classes until the rule is satisfied is recommended. This topic has recently been subject to vigorous discussion among the psychologists [17], [18]. The numbers 10 and 5 appear to have been arbitrarily chosen. A few investigations throw some light on the appropriateness of the rule. The approach has been to examine the exact distribution of $X^2$, when some or all expectations are small, either by mathematical methods or from sampling experiments.

The investigations are scanty and narrow in scope, as is to be expected since work of this type is time-consuming. Thus the recommendations given below may require modification when new evidence becomes available.

To digress for a moment, the problem of investigating the behavior of $X^2$ when expectations are small is an example of a whole class of problems that are relevant to applied statistics. In applications it is an everyday occurrence to use the results of a body of theory in situations where we know, or strongly suspect, that some of the assumptions in the theory are invalid. Thus the literature contains investigations of the $t$-distribution when the parent population is nonnormal, and of the performance of linear regression estimates when the regression in the population is actually nonlinear. Fortunately for applications, the results of theory sometimes remain substantially true even when some assumptions fail to hold. This fact tends to make statistics a more confusing subject than pure mathematics, in which a result is usually either right or wrong.

In any problem of this kind, it is important to define what is meant by saying the results remain "substantially true." I stress this point because a reader who becomes interested in a specific problem and tries to summarize the available knowledge may encounter considerable difficulty. Definitions vary from writer to writer and are sometimes entirely subjective, so that the researches may be presented in a form which baffles any attempt to apply a uniform definition. This remark is not intended as a criticism of work on the $X^2$ problem, where the task of summarizing is comparatively easy. However, I believe that the usefulness of this kind of research would be enhanced by careful attention to the questions: (i) how are we going to measure the disturbance caused by a failure in assumptions, and (ii) when are we going to call this disturbance "serious."

In the present instance, my criterion is to compare the exact $P$ and the $P$ from the $\chi^2$ table, when the null hypothesis is true, in the region in which the tabular $P$ lies between 0.05 and 0.01. This criterion is not ideal, but it does appraise the performance of the tabular approximation in the borderline region between statistical significance and nonsignificance. A disturbance is regarded as unimportant if when the $P$ is 0.05 in the $\chi^2$ table, the exact $P$ lies between 0.04 and 0.06, and if when the tabular $P$ is 0.01, the exact $P$ lies between 0.007 and

0.015. These limits are, of course, arbitrary; some would be content with less conservative limits.

The results suggest that four cases need to be considered separately.

(i) *Goodness of fit tests of bell-shaped curves such as the normal distribution.* The distinguishing feature of this case is that usually only one or two expectations at the tails are small, the others being above the conventional limits of 5 or 10. Cochran [19] has shown that there is little disturbance to the 5% level when a *single* expectation is as low as $\frac{1}{2}$.

This is also true for the 1% level if the number of degrees of freedom in $X^2$ exceeds 6. Two expectations as low as 1 may be allowed with negligible disturbance to the 5% level. Since the discrepancy between an observed and a postulated distribution is often most apparent at the tails, the sensitivity of the $X^2$ test is likely to be decreased by an overdose of pooling at the tails. Thus considerations of the power of the test urge us to use cells with as small expectations as we dare from distributional considerations. The inflexible use of minimum expectations of 5 or 10 may be harmful.

(ii) $2 \times 2$ *contingency tables.* This case is the most thoroughly worked out and can be regarded as solved for practical purposes. Fisher [15] has given the method of obtaining an exact solution, which is not too laborious in samples up to size 30. Tables such as Mainland's [20] give the probability levels of the exact distribution for two samples each of size up to $n = 40$, and Yates' table [21] gives almost exact tests based on $X^2$ after correction for continuity.

(iii) *Tests in which all expectations may be small.* This case occurs from time to time, for example, in genetical studies in which a Mendelian ratio is being compared over a number of small families. Results by Neyman and Pearson [22], Cochran [23] and Sukhatme [24] suggest tentatively that the tabular $\chi^2$ is tolerably accurate provided that all expectations are at least 2.

With very scanty data, there is one danger—that only a few different values of $X^2$ are possible, so that the effects of discontinuity become noticeable. For example, consider the $2 \times 4$ contingency table with marginal totals shown below. All expectations are exactly 2.

| | | | | | |
|---|---|---|---|---|---|
| | | | | | 8 |
| | | | | | 8 |
| | 4 | 4 | 4 | 4 | 16 |

If we construct all tables which satisfy these marginal totals, only seven different values of $X^2$ are found. The exact distribution of $X^2$ and the $\chi^2$ approximation (with 3 degrees of freedom) are shown in Table 2. The agreement is not good, the tabular $P$'s being fairly consistently too low.

With such a small number of values of $X^2$, a correction for continuity comes to mind. To apply this for $X^2 = 2$, we read the $\chi^2$ table at $\chi^2 = 1$, this being

half way between 2 and 0 (the next largest value of $X^2$). The corrected $P$'s show a considerable improvement in fit.

In practice, a small table of this kind can be handled by computing the exact distribution of $X^2$ in cases of doubt about the adequacy of the $\chi^2$ approximation. For more complex contingency tables, a systematic method of computation has been given by Freeman and Halton [25].

TABLE 2

*Exact distribution of $X^2$ for a 2 × 4 contingency table. $P\{X^2 \geqq X_0^2\}$*

| $X_0^2$ | Exact | $\chi^2$ Table | Corrected |
|---|---|---|---|
| 0 | 1.000 | 1.000 | 1.000 |
| 2 | .899 | .572 | .801 |
| 4 | .362 | .261 | .391 |
| 6 | .243 | .112 | .172 |
| 8 | .064 | .046 | .072 |
| 10 | .030 | .019 | .029 |
| 16 | .0005 | .0011 | .004 |

(iv) *Tests in which all expectations are small and $X^2$ has many degrees of freedom (say >60).* Examples occur in genetical research. The data are presented in, say, a 2 × 200 contingency table, with all 400 expectations small.

In this case, the exact distributions of $X^2$ and $\chi^2$ are both approximately normal, since the degrees of freedom are large. However, the two distributions have different variances, and the normal approximation to the exact distribution is sometimes quite different from the normal approximation to $\chi^2$.

This problem has been studied by Haldane [26], [27] who has worked out the exact mean and variance of $X^2$ for several types of data. His results are given below for the two cases that are perhaps most common.

(a) We have $g$ groups, each containing $s$ individuals, classified into one of two classes. The null hypothesis specifies a *known* constant probability $p$ that any individual falls into the first class. If $x_i$ individuals fall into this class in the $i$th group, and if we wish to test against the alternative that $p$ varies from class to class, a familiar extension of the $X^2$ test is to calculate

$$X^2 = \sum_{i=1}^{g} \frac{(x_i - sp)^2}{spq}$$

with $g$ degrees of freedom. Haldane shows that

$$E(X^2) = g,$$

$$V(X^2) = 2g\left(1 + \frac{1 - 6pq}{2spq}\right).$$

(b) Same data, but $X^2$ computed as for a $2 \times g$ contingency table, since $p$ is unknown.

| | | |
|---|---|---|
| | | $n_1$ |
| | | $n_2$ |
| $s$ | $s \quad \cdots \quad s$ | $n$ |

Then

$$E(X^2) = \frac{(g-1)n}{n-1},$$

$$V(X^2) = \frac{2(g-1)n^3(n-g)}{(n-1)^2(n-2)(n-3)}\left(1 - \frac{(n-1)}{n_1 n_2}\right).$$

To take an extreme case, suppose $s = 2$, $g = 160$, $n = 320$, $n_1 = 64$, $n_2 = 256$. The mean and variance of $X^2$ are 159.5 and 159.4 respectively, whereas $\chi^2$ has mean 159 and variance 318, twice as large. The normal approximation to $X^2$ is satisfactory, but $\chi^2$ gives very poor agreement [23]. In practice, such data may be dealt with by use of the normal approximation to the exact distribution, using Haldane's expressions for the mean and variance.

The question remains: where does case (iii) shade into case (iv)? The available data suggest that case (iii) may apply when $X^2$ has less than 15 degrees of freedom, while case (iv) may hold if $X^2$ has more than 60 degrees of freedom. The intervening gap needs investigation.

**8. The correction for continuity.** The exact distribution of $X^2$ is always discontinuous. When all expectations are small, the number of distinct values of $X^2$ may be very limited. In such cases the $\chi^2$ table may give a poor approximation to the exact $P\{X^2 \geqq X_0^2\}$, mainly because the area of a continuous curve is used to approximate the sum of a small number of discrete probabilities. The correction for continuity, introduced by Yates [28], is an attempt to remove this source of error. It amounts to reading the $\chi^2$ table, not at the point $X_0^2$, but at a point halfway between $X_0^2$ and the value of $X^2$ immediately below $X_0^2$ in the discrete series of values.

In practice, the correction is seldom needed except when $X^2$ has 1 degree of freedom, as when testing a single binomial ratio or a $2 \times 2$ contingency table. In the $2 \times 2$ table there are various ways of computing the correction, depending on how one likes to compute $X^2$. My own preference is to find the difference $d$ between $x_i$ and $m_i$, which is the same, apart from sign, in all four cells. The absolute value of $d$ is reduced by $\frac{1}{2}$, and $X^2$ is computed as

$$X^2 = (|d| - \tfrac{1}{2})^2 \sum \frac{1}{m_i}.$$

Note that in the $2 \times 2$ table it is $X$ that is corrected for continuity, not $X^2$, since the successive values of $d$ differ by unity.

When $X^2$ has 1 degree of freedom, a good rule is to apply the correction whenever it produces any appreciable difference in the significance probability. The correction has a tendency to over-correct, changing the tabular $P$ from too small to too large, but is usually an improvement.

If a number of $X^2$ values, each with 1 degree of freedom, are added to form a total $X^2$, the individual $X^2$ values should *not* be corrected for continuity, because the over-correction mounts up in a disconcerting manner [19]. After it has been obtained, the total $X^2$ is corrected, if this is necessary, by the method given in the following paragraph.

Compute the next largest value of $X^2$ which the structure of the data permits. Read the $\chi^2$ table at a point halfway between this value and the observed $X^2$. This procedure was illustrated for a $2 \times 4$ table in the preceding section. Sometimes the next largest value of $X^2$ is not immediately obvious and trial and error is required to find it.

**9. The construction of classes.** When $X^2$ is used to test the hypothesis that the observations follow a *continuous* frequency distribution, the first step is to group the observations into classes. Both the number of classes and the division points between classes are at the disposal of the investigator, and the choices that he makes will affect the sensitivity of the test. I believe that the common practice is to have a moderate number of classes, say between 10 and 25, and to make the class intervals equal. Although information about the best rule for constructing the classes is still meager, the recommendations of those who have looked into this problem are contrary to current practice.

With regard to class intervals, Mann and Wald [29] and Gumbel [30] suggest that these be chosen so that the *expected* number is the same ($= n/k$) in all classes. These authors do not claim that this will increase the power of the test, but merely suggest that it is likely to be a good procedure. Gumbel points out that if this method is used in conjunction with a rule for choosing the value of $k$, much of the arbitrariness that accompanies the construction of class intervals is removed. Under this method, the computational steps are first to estimate the constants (mean, s.d., etc.) which determine the fitted curve, then find the class boundaries which give equal expectations in each class, and finally count the *observed* numbers $x_i$ which fall in the respective classes. The value of $X^2$ is given as

$$X^2 = \frac{k}{n} \sum x_i^2 - n.$$

The paper by Mann and Wald deals with the choice of the number of classes, $k$. The null hypothesis is assumed to specify the distribution completely, and $n$ is assumed large enough so that the limiting $\chi^2$ distribution is applicable.

Some criterion is required to define what is meant by a "best" value of $k$.

It seems natural to try to maximize some property of the power function of the test. The criterion set up by Mann and Wald is a little complex to describe, but this stems from the complexity of the problem.

They define the distance $\Delta$ between the null distribution and any alternative distribution as the maximum difference between the heights of the two cumulative distribution functions. It becomes evident, after some examination of the problem, that there is no hope of choosing $k$ so as to maximize the power function of $X^2$ at all points along its course. They decide to concentrate on maximizing the power at about the point where the power is $\frac{1}{2}$. This is an arbitrary but reasonable choice. The two principal properties possessed by their "best" $k$ are as follows.

(i) For a value of $\Delta$ which they determine, the power of the $X^2$ test is at least $\frac{1}{2}$ for all alternative distributions whose distance from the null distribution is at least $\Delta$. This value $\Delta$ is a simple function of sample size and, as would be expected, it decreases steadily with increasing sample size.

(ii) If any $k$ other than the "best" is chosen, the power of $X^2$ is less than $\frac{1}{2}$ for at least one alternative whose distance from the null distribution exceeds $\Delta$.

The best $k$ is given by the formula

$$k = 4\left[\frac{2(n-1)^2}{c^2}\right]^{\frac{1}{5}},$$

where

$$\frac{1}{\sqrt{2\pi}}\int_c^\infty e^{-(x^2/2)}\,dx = \alpha,$$

where $\alpha$ is the significance level. Thus $c = 1.64$ for a test at the 5% level.

The optimum values of $k$ are substantially higher than those customary in practice. For a test at the 5% level, $k$ rises slowly from 31 at $n = 200$ to 78 at $n = 2{,}000$.

TABLE 3

*Optimum number of classes (Mann and Wald)*

| $n$ | 200 | 400 | 600 | 800 | 1,000 | 1,500 | 2,000 |
|---|---|---|---|---|---|---|---|
| $k$ | 31 | 41 | 48 | 54 | 59 | 70 | 78 |

A good exposition and critique of the Mann-Wald paper has been given by Williams [31]. The Mann-Wald method is more tedious to compute than the usual procedure, partly because of the increased number of classes and partly because of the fitting with equal *expected* numbers. Williams shows, however, that the optimum is a broad one, and that the value of $k$ in Table 3 can probably be halved with little loss in sensitivity.

The Mann-Wald paper, although an able performance in a difficult field, is far from a complete investigation of the optimum number of classes. Such an

investigation is unlikely to be forthcoming soon. What is the user of the $X^2$ test to conclude from their results? My own reaction has been to put more computational work into $X^2$ tests of continuous distributions, by increasing the number of intervals and using unequal lengths of interval where this is necessary in order to avoid classes with high expected numbers. For sample sizes between 200 and 1000, their recommended expected numbers per class in Table 3 range from 6 to 16. With Williams' modification, the range is from 12 to 32. This does suggest that there is an appreciable loss of power if classes with expectations of more than 50 are commonly used.

**10. Summary recommendations.** The following is an attempt to set down in brief form the recommendations about the computation of $X^2$ which flow from the discussion in this part and from practical experience. The recommendations are not as explicit as I should like. They can, I believe, be made more explicit, but this requires detailed study that goes beyond the scope of the present paper. The total number of observations is $n$.

*I. Attribute data.* The data come to us in grouped form. Pooling of classes is considered undesirable because of loss of power.

(a) *The* $2 \times 2$ *table.* Use Fisher's exact test (i) if $n < 20$, (ii) if $20 < n < 40$ and the smallest expectation is less than 5. Mainland's tables [20] are helpful in all such cases. If $n > 40$, use $X^2$, corrected for continuity if the smallest expectation is less than 500.

(b) *Tables with degrees of freedom between* 2 *and* 60 *and all expectations less than* 5. If $n$ is so small that Fisher's exact test can be computed without excessive labor, use this. Otherwise use $X^2$, considering whether this needs correction for continuity by finding the next largest value of $X^2$.

(c) *Tables with degrees of freedom greater than* 60 *and all expectations less than* 5. Try to obtain the exact mean and variance of $X^2$ and use the normal approximation to the exact distribution.

(d) *Tables with more than* 1 *degree of freedom and some expectations greater than* 5. Use $X^2$ without correction for continuity.

*II. Continuous data.* The data must first be grouped. Use enough cells to keep the expectations down to the levels recommended by Williams (12 per cell for $n = 200$, 20 per cell for $n = 400$, 30 per cell for $n = 1{,}000$). At the tails, pool (if necessary) so that the minimum expectation is 1.

## Part III. Utility of the Test

**11. Criticisms and limitations of the test.** A competent appraisal of the utility of the $X^2$ test would require a sampling survey of scientists and others who try to draw conclusions from data. In such a survey the object would be to discover how frequently these workers have occasion to use a $X^2$ test, and to what extent the application of the test really seems to help them. In fact, such a survey, directed at the use of statistical techniques in general and not merely at the $X^2$ test, might be very illuminating to statisticians if it could be carried out despite the obvious difficulties. Statisticians are, I think, rather quick to jump to con-

clusions about the kinds of problems which scientists in other fields are supposed to face, and about their presumed uses and misuses of statistical methods and ideas.

In the absence of survey data of this kind, the statistician can give only a personal opinion, based on such contacts as he has had with the users of the $X^2$ test. I will content myself with the cautious statement that since the construction of hypotheses and their continued modification or rejection in the light of new data is one of the standard tools of science, some kind of test of the agreement between theory and data must often be useful. The experiences of Airy and Merriman illustrate the uncomfortable position in which the scientist is placed when he has to state, without the benefit of such a test, whether his observations are in accordance with the predictions of some theory.

On the other hand, a reading of the literature reveals the opinion, expressed by several writers, that the $X^2$ test is of restricted usefulness. The reasons for this critical verdict seem to be diverse. Some of the criticism is directed at the $X^2$ test itself, but some seems to apply to composite, or "general purpose" tests of significance as a whole, and some to *all* tests of significance.

Considering first the criticisms of $X^2$ itself, the name "goodness of fit" is misleading, because the power of the test to detect an underlying disagreement between theory and data is controlled largely by the size of the sample. With a small sample, an alternative hypothesis which departs violently from the null hypothesis may still have a small probability of yielding a significant value of $X^2$. In a very large sample, small and unimportant departures from the null hypothesis are almost certain to be detected. Consequently, when $X^2$ is nonsignificant, the amount by which the null hypothesis has been strengthened depends mainly on the size of sample. This is one of the principal reasons for such misuse of the test as exists. Authors sometimes write as if the validity of their null hypothesis has been greatly strengthened, if not definitely established, by a goodness of fit test made on very scanty data.

Secondly, as Gumbel has pointed out, the $X^2$ test for a continuous frequency distribution is not unique, because of the freedom of choice of number of intervals and end-points of the intervals. Although this is an argument for more standardization in the application of the test, the objection perhaps is minor rather than major. At least, statisticians have not seen any overwhelming advantage in having just one test of a given null hypothesis against a given alternative. In recent years there has been active research in the development of quick, though inefficient, tests for problems in which satisfactory, but less speedy, tests already exist. The tests will give different values of $P$ from the same data, but no serious objections to this situation seem to have been noticed.

There are two available substitute tests which resemble the $X^2$ test in that they are not directed against any specific alternative. One is the $\omega^2$ test (Section 13). This was constructed in order to avoid the grouping of continuous data that is necessary with $X^2$. The other, for data that are in grouped form, is the likelihood ratio test against a completely general alternative hypothesis (Section 14).

One limitation of $X^2$, or of any nonspecific test, is that when the alternative

hypotheses can be fairly clearly defined, we may hope to obtain another test, directed against these alternatives, that will be more powerful than $X^2$. An example is Neyman's "smooth" tests (Section 15). These were constructed to detect alternative hypotheses that depart from the null hypothesis in some continuous or smooth fashion. Like $X^2$, the smooth tests are still general, since they do not demand detailed knowledge of the nature of the alternatives. Further down the scale there is a variety of supplementary tests to $X^2$.

Finally, the $X^2$ test is sometimes used when what is needed is not a test of significance of the usual type. There are numerous occasions when the null hypothesis is not expected to be exactly true, but at best approximately true. The argument against $X^2$ in this situation has been developed amusingly by Berkson [32]. He writes "I make the following dogmatic statement, referring for illustration to the normal curve: 'If the normal curve is fitted to a body of data representing any real observations whatever of quantities in the physical world, then if the number of observations is extremely large—for instance, on the order of 200,000—the chi-square $P$ will be small beyond any usual limit of significance.'"

If this statement is granted—and counter-evidence, to put it mildly, is not abundant—then Berkson proceeds to the Socratic conclusion. What is the point of applying a $X^2$ test to a moderate or small sample if we already know that a large sample would show $P$ highly significant?

In his original paper, Karl Pearson was aware of this issue, and did not seem to feel uncomfortable about it. He writes, "Nor again does it appear to follow that if the number be largely increased the same curve will still be a good fit. Roughly the $\chi^2$'s of two samples appear to vary for the same grouping as their total contents. Hence if a curve be a good fit for a large sample, it will be good for a small one, but the converse is not true, and a larger sample may show that our theoretical frequency gives only an approximate law for samples of a certain size. In practice we must attempt to obtain a good fitting frequency for such groupings as are customary or utile. To ascertain the ultimate law of distribution of a population for any groupings, however small, seems a counsel of perfection." Although it is hazardous to try to read another man's mind, his attitude was apparently the defensible one that any theory is at best approximately true, but nevertheless, if we are going to reject a theory, we do so because it does not fit the data that we have, not because it would not fit a much larger sample of data that we do not have.

Nevertheless, I would agree with Berkson that in this situation an ordinary test of significance is not very useful. It is more difficult to say just what we do want. One attack would be to reformulate the null hypothesis so that, instead of testing whether a binomial $p$ equals $p_0$, we try to construct a test of the null hypothesis that $p$ lies in the specified range $p_0$, $p_1$.

As an alternative approach, fiducial or confidence limits seem to be helpful. Suppose that these limits are set up for the difference between two percentages in a $2 \times 2$ contingency table, the ordinary null hypothesis being that the true difference is zero. If the two limits are far from zero, then even when $X^2$ is non-

significant we are warned that the data do not establish the null hypothesis as approximately true. If the limits are both near zero, on the other hand, we may be able to conclude that the null hypothesis, although presumably not exactly true, is close enough to the correct hypothesis for all practical purposes.

In testing goodness of fit of a frequency distribution, the extension of this approach is Kolmogorov's method [33] for constructing confidence bounds for the cumulative frequency distribution, given a sample.

To summarize, the $X^2$ test is helpful primarily in the exploratory stages of an investigation, when there is no very clear knowledge of the alternative hypotheses. It is well to remember that the size of sample determines whether the test really is a severe test of the null hypothesis.

**12. Interpretation of high *P*'s.** The question of the interpretation to be placed on very high $P$'s, say those greater than 0.99, has been raised from time to time. In the few instances of this kind that have come my way, my practice has been to give the data further scrutiny before regarding the result as evidence in favor of the null hypothesis.

Events have justified this practice. In nearly every instance, something wrong was discovered, most frequently a numerical mistake or an error in the formula used to compute $X^2$. In one set of data assembled by geneticists, a whole group of $X^2$ showed $P$'s of the order of 0.999. The reason was that these $X^2$ values had been obtained by adding a large number of $X^2$ values, each with 1 degree of freedom, and all the original (1 d.f.) $X^2$ had been corrected for continuity. The over-correction which is a feature of this device had piled up to such an extent that their total $X^2$'s were much smaller than those following the $\chi^2$ distribution. In another case, after discussion with the assistants of the scientist in charge, I surmised that the observations had been influenced by the anticipations of the scientist. Fisher [34] has raised a similar speculation with respect to some of Mendel's results, without any suggestion of improper scientific conduct on the part of Mendel.

## Part IV. Tests Which Are Competitive to $X^2$

**13. The $\omega^2$ test.** Alternatives that have been proposed to the $X^2$ test are of two kinds. Several of the tests, like $X^2$, are "general" tests. Then there is a battery of supplementary tests that are intended for situations where the alternative hypothesis is more definitely known.

The *general* substitute tests that have been proposed have not given $X^2$ very serious competition. This is understandable because of the long history of $X^2$ and of its inclusion as standard doctrine in most elementary courses, and because some of the substitute tests are limited in the type of hypothesis with which they can cope. Moreover, despite the weaknesses of the $X^2$ test discussed in Part III, the advantages of the alternative tests have not yet been clearly enough demonstrated to win many converts. Consequently, Part IV contains only a brief and rather noncommital introduction to these tests.

The first general test, developed by Cramér [35], von Mises [36] and Smirnov [37], was constructed mainly for use with small samples. The null hypothesis completely specifies the frequency distribution followed by the observations. Unlike $X^2$, the $\omega^2$ test requires no grouping of the observations, an obvious advantage with small samples. The test is based on a comparison of the cumulative frequency function $F(x)$ specified by the null hypothesis with an estimate of the cumulative frequency made from the sample. This estimate, $F^*(x)$, is simply $r/n$, where $r$ is the number of observations in the sample which are $\leqq x$. The test criterion proposed is the Stieltjes integral

$$\omega^2 = \int_{-\infty}^{\infty} [F(x) - F^*(x)]^2 \, dF(x).$$

If $F(x)$ is continuous, this may be shown to satisfy

$$\omega^2 = \frac{1}{12n^2} + \frac{1}{n} \sum_{r=1}^{n} \left[ F(x_r) - \frac{2r-1}{2n} \right]^2,$$

where the values $x_1, x_2, \cdots, x_n$ are now arranged in increasing order.

The mean and variance of $\omega^2$ are known, and also its limiting distribution (which is nonnormal) as $n \to \infty$. A table of this distribution by Darling and Anderson [38] has appeared recently. Practical use of this test is restricted by the condition that $F(x)$ must be known and by lack of information about the small-sample distribution of $\omega^2$.

**14. The likelihood ratio test.** If the data are presented in grouped form, and if the alternative hypothesis is completely general, it is known that in large samples the $X^2$ test and the likelihood ratio test become equivalent [9]. We start from the usual multinomial

$$Pr = \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}.$$

The likelihood on the null hypothesis is found from $p_i = m_i/n$, where $m_i$ are the expectations estimated by maximum likelihood (unless they are explicitly given). The likelihood on the unrestricted alternative is found from $p_i = x_i/n$. Hence the likelihood ratio becomes

$$\left(\frac{x_1}{m_1}\right)^{x_1} \left(\frac{x_2}{m_2}\right)^{x_2} \cdots \left(\frac{x_k}{m_k}\right)^{x_k}.$$

Its logarithm is

$$L = \sum_{i=1}^{k} x_i \log\left(\frac{x_i}{m_i}\right) = \sum_{i=1}^{k} x_i \log\left\{1 + \frac{x_i - m_i}{m_i}\right\}.$$

When this is expanded in a power series in the $(x_i - m_i)$, the leading term is $X^2$ for the maximum likelihood estimates of the parameters.

In view of the equivalence of the two criteria in large samples, there seems no advantage, except one of taste or convenience, in one test over the other.

For small samples, the suggestion has been made from time to time that the likelihood ratio is to be preferred. Examples worked by both tests have been presented and discussed by Neyman and Pearson [22] and Fisher [39]. Since users of statistical methods naturally do not wish to learn more tests than are necessary, a movement to replace $X^2$ by the likelihood ratio seems unlikely to gather momentum unless some definite advantages can be shown to follow. The advantage in computing time is at most small, but there may be an increase in power. The striking way in which many different configurations of the data turn out to give exactly the same value of $X^2$ in small samples suggests an element of coarseness in the $X^2$ test. This coalescence happens to a much reduced extent with the likelihood ratio. However, not enough data about relative power has accumulated to permit a verdict on this issue.

**15. Neyman's smooth tests.** As in the $\omega^2$ test, Neyman [40] postulates that the cumulative frequency $F(x)$ (assumed continuous) is known exactly from the null hypothesis. The first step is to replace the observations $x_i$ by the familiar "probability integral" transforms $y_i$, where

$$y_i = F(x_i).$$

If the null hypothesis is correct, the variates $y_i$ follow a rectangular distribution in the interval (0, 1). The problem, therefore, reduces to that of finding a test for this transformed hypothesis.

Neyman points out that the conceivable alternatives to the null hypothesis fall into two broad classes. The first class, of "smooth" alternatives, contains frequency functions which are continuous and which depart in some gradual and regular manner from the null hypothesis. The second class contains all other alternatives, whose deviation from the null hypothesis is in some respects erratic or discontinuous. The $X^2$ test is not directed specifically at either class, and is to some extent effective against both types of departure from the null hypothesis. Neyman's object is to develop tests sensitive to the first class of alternatives.

If a "smooth" alternative holds, the transforms $y_i$ will no longer follow a rectangular distribution, but will presumably follow a continuous distribution with a limited number of maxima and minima. The proposal is, therefore, to test the $y_i$ for polynomial trends, on the assumption that a polynomial of fairly small degree will satisfactorily represent the smooth alternative. The computations involved and a discussion of the appropriate degree of the polynomial are presented in [40].

## Part V. Tests Which Are Supplementary to $X^2$

**16. A supplementary test based on runs.** $X^2$ takes no account of the succession of + and − signs in the deviations $(x_i - m_i)$ between observations and expectations. When a smooth alternative holds, it seems likely that the succession of signs will exhibit some systematic features, such as a long run of +'s followed

by a run of $-$'s, and this has often been observed in applications of $X^2$. David [41] has adapted the now familiar "run" test as a supplementary test to $X^2$. In the run test, we count the number of runs and refer to a table which shows the significance levels of this quantity, given the total numbers of $+$'s and $-$'s in the series. David has shown that the limiting distribution of the number of runs is independent of that of $X^2$.

The run tests will, of course, be most effective for alternatives which produce few runs, such as a shift in the mean of the distribution. In the reference cited, the test is restricted to the case where the null hypothesis completely specifies the distribution; David states that a test has been developed for the case in which some parameters must be estimated.

**17. Tests based on low moments.** When the null hypothesis postulates that the observations follow a normal, binomial or Poisson distribution, an alternative to $X^2$ that is in fairly common use is to compare the lower moments of the theoretical distribution with estimated moments from the sample. With the normal distribution, the actual values of the mean and variance are rarely given by the null hypothesis, so that a comparison of these moments is not usually possible. Tests of skewness, derived from the third moment, and of kurtosis, derived from the fourth moment, can be made [15].

In the binomial distribution, if $p$ is specified we can compare the sample mean and variance with the theoretical mean and variance. If $p$ is not specified, we can compare variances. Suppose that we have $g$ series of trials, and that each series contains $s$ trials. The number of successes (out of $s$) in the $i$th trial is $x_i$. For a test of the mean proportion of successes, we regard

$$\frac{\left(\frac{\bar{x}}{s} - p\right)}{\sqrt{\frac{pq}{gs}}}$$

as a normal deviate.

**18. Dispersion tests.** Turning to the variance of $x_i$, if $p$ is specified the estimated variance is $\sum (x_i - sp)^2/g$, while the theoretical variance is $spq$. An appropriate test criterion for the variance is, therefore,

$$\frac{\sum (x_i - sp)^2}{gspq}.$$

If $p$ must be estimated from the data, either because it is unspecified or because the sample estimate disagrees with the postulated $p$, the test criterion becomes

$$\frac{\sum (x_i - \bar{x})^2}{gs\left(\frac{\bar{x}}{s}\right)\left(1 - \frac{\bar{x}}{s}\right)} = \frac{s\sum (x_i - \bar{x})^2}{g\bar{x}(s - \bar{x})}.$$

As is well known, the related quantities

$$\frac{\sum (x_i - sp)^2}{spq} \quad \text{and} \quad \frac{s \sum (x_i - \bar{x})^2}{\bar{x}(s - \bar{x})} \tag{13}$$

are distributed approximately as $\chi^2$ with $g$ and $(g - 1)$ degrees of freedom, respectively, when the null hypothesis is true.

The variance test can also be made when the number of trials $s_i$ varies from series to series. The test criterion becomes

$$\frac{\sum s_i \left(\frac{x_i}{s_i} - \frac{\bar{x}}{\bar{s}}\right)^2}{\frac{\bar{x}}{\bar{s}}\left(1 - \frac{\bar{x}}{\bar{s}}\right)} = \frac{\bar{s}^2 \left\{\sum \frac{x_i^2}{s_i} - \frac{(\sum x_i)^2}{\sum s_i}\right\}}{\bar{x}(\bar{s} - \bar{x})}.$$

This test criterion can be shown to be identical with $X^2$ as calculated for the $2 \times g$ contingency table.

| | | |
|---|---|---|
| | $x_1 \quad x_2 \quad \cdots \quad x_g$<br>$s_1 - x_1 \quad s_2 - x_2 \quad \cdots \quad s_g - x_g$ | $\sum x_i$<br>$\sum s_i - \sum x_i$ |
| | $s_1 \quad s_2 \quad \cdots \quad s_g$ | $\sum s_i$ |

It is important to distinguish clearly between this variance $X^2$ test, which is sometimes called a *dispersion* test, and the ordinary goodness of fit $X^2$ test. Suppose that we have 200 families each of size 4, and that every individual belongs to one of two classes $A$ and $a$. The null hypothesis states that the probability $p$ of an $A$ is constant for all families. We may tally the numbers of families that have 0, 1, 2, 3, 4 $A$'s, respectively, and test this frequency distribution against the binomial $(q + p)^4$. This is the ordinary goodness of fit test, which has 3 degrees of freedom if $p$ is unknown. The dispersion test, computed by formula (13), compares the observed variance of this frequency distribution with the theoretical binomial variance.

The dispersion tests frequently prove more sensitive than $X^2$ when the binomial null hypothesis fails because the probability of an $A$ varies from one family to another. The notion of a measure of dispersion of this kind is due to Lexis and antedates the goodness of fit test.

**19. Subdivision of $X^2$ into components.** In the analysis of variance, the subdivision of a sum of squares into single components, or "single degrees of freedom," is a familiar device. If variates $y_i$ are normally and independently distributed with mean 0 and variance $\sigma^2$ on some null hypothesis, these components are obtained by any linear transformation of the form

$$z_i = \sum_{j=1}^{k} l_{ij} y_j \qquad (i = 1, 2, \cdots k)$$

where

$$\sum_j l_{ij} l_{hj} = \begin{cases} 0, & i \neq h, \\ 1, & i = h. \end{cases} \tag{14}$$

All $z_i$ are normally and independently distributed with mean 0 and variance $\sigma^2$. This transformation enables us to select those $z_i$ that are likely to be sensitive to a particular alternative hypothesis. Often only one or two of the $z_i$ are examined, because it is hard to imagine any feasible alternative that would make the other $z$'s large. Thus the device replaces a "sum of squares" test by a few more specialized tests.

The corresponding subdivision of $X^2$ is easily obtained from Fisher's device of regarding the observed numbers $x_i$ in the cells as following independent Poisson distributions, subject to a single linear restriction. Thus when all expectations are large, we may take

$$y_i = \frac{x_i - m_i}{\sqrt{m_i}}$$

as the set of unit normal deviates. Since these are subject to the linear restriction

$$\sum_{j=1}^{k} (x_j - m_j) = \sum_{j=1}^{k} \sqrt{m_j}\, y_j = 0,$$

we must take

$$Z_1 = \sum \sqrt{m_j}\, y_j / n.$$

Let the remaining $Z_i$ ($i = 2, 3, \cdots k$) be

$$Z_i = \sum_j l'_{ij} x_j = \sum l'_{ij} \sqrt{m_j}\, y_j + \sum l'_{ij} m_j.$$

If these are to have means zero, we must have

$$\sum_j l'_{ij} m_j = 0. \tag{15}$$

Note that this relation makes all the remaining $Z_i$ orthogonal with $Z_1$. Since $l_{ij} = l'_{ij} \sqrt{m_j}$, equations (14) become

$$\sum_j l'_{ij} l'_{hj} m_j = \begin{cases} 0 & i \neq h, \\ 1 & i = h. \end{cases} \tag{16}$$

Any set of $Z_i$ whose coefficients satisfy equations (15) and (16) provide a breakdown of $X^2$ into $(k - 1)$ single components. Then as an algebraic identity,

$$X^2 = \sum_{i=1}^{k} \frac{(x_i - np_i)^2}{np_i} = \sum_{i=1}^{k-1} Z_i^2.$$

As $n$ increases, the individual terms on the right become in the limit independently distributed as $\chi^2$ with 1 degree of freedom. In genetic analysis,

where simple interpretations can be attached to the $Z_i$, this tool has proved useful [15].

The application of this breakdown to contingency tables, which requires care, has been elucidated by Lancaster [42] and Irwin [43]. In an $r \times c$ contingency table, $X^2$ can be partitioned into $(r-1)(c-1)$ single components. Each of these represents the usual $X^2$ for a $2 \times 2$ table which is formed by amalgamation of cells in the original table. This breakdown is illustrated below for a $3 \times 3$ table.

*Original table*

| | | | |
|---|---|---|---|
| $x_{11}$ | $x_{12}$ | $x_{13}$ | $r_1$ |
| $x_{21}$ | $x_{22}$ | $x_{23}$ | $r_2$ |
| $x_{31}$ | $x_{32}$ | $x_{33}$ | $r_3$ |
| $c_1$ | $c_2$ | $c_3$ | $n$ |

*Components*

| | | | | | |
|---|---|---|---|---|---|
| $x_{11}$ | $x_{12}$ | $r_{12}$ | $r_{12}$ | $x_{13}$ | $r_1$ |
| $x_{21}$ | $x_{22}$ | $r_{22}$ | $r_{22}$ | $x_{23}$ | $r_2$ |
| $c_{21}$ | $c_{22}$ | $n_{22}$ | $n_{22}$ | $c_{23}$ | $n_{23}$ |
| $c_{21}$ | $c_{22}$ | $n_{22}$ | $n_{22}$ | $c_{23}$ | $n_{23}$ |
| $x_{31}$ | $x_{32}$ | $r_{32}$ | $r_{32}$ | $x_{33}$ | $r_3$ |
| $c_1$ | $c_2$ | $n_{32}$ | $n_{32}$ | $c_3$ | $n$ |

If the $X^2$ are calculated in the usual way for each $2 \times 2$ table, the partition is only approximate, in that in finite samples the individual $X^2$ do not add up to the total $X^2$ for the $3 \times 3$ table. The authors show how to obtain a partition which adds up exactly, that is, which satisfies the sets of equations (15) and (16). It appears that the approximate partition is adequate for most tests of significance; in fact, it has not been shown that the additive partition is really preferable to the approximate partition in small samples.

Another application of the breakdown of $X^2$ is to contingency tables in which numerical scores can be attached to one or both of the classifications. Yates [44] shows how to isolate and compare the regressions of the observations on these scores.

In conclusion, the testing of individual components of $X^2$ is analogous to the use of a set of independent $t$-tests instead of, or in addition to, an $F$-test in the analysis of variance.

I wish to thank T. W. Anderson, E. L. Lehmann and J. W. Tukey for many helpful suggestions.

## REFERENCES


[1] KARL PEARSON, "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," *Philos. Mag. Series 5*, Vol. 50 (1900), pp. 157–172.

[2] H. HOTELLING, "The consistency and ultimate distribution of optimum statistics," *Trans. Am. Math. Soc.*, Vol. 32 (1930), pp. 847–859.

[3] R. A. FISHER, "On the interpretation of chi square from contingency tables, and the calculation of $P$," *Jour. Roy. Stat. Soc.*, Vol. 85 (1922), pp. 87–94.

[4] M. GREENWOOD AND G. U. YULE, "The statistics of anti-typhoid and anti-cholera inoculations and the interpretation of such statistics in general," *Proc. Roy. Soc. Med.*, Vol. 8 (1915), pp. 113–190.

[5] R. A. FISHER, "The conditions under which chi square measures the discrepancy between observation and hypothesis," *Jour. Roy. Stat. Soc.*, Vol. 87 (1924), pp. 442–450.

[6] G. U. YULE, "On the application of the $\chi^2$ method to association and contingency tables, with experimental illustrations," *Jour. Roy. Stat. Soc.*, Vol. 87 (1922), pp. 76–82.

[7] J. BROWNLEE, "Some experiments to test the theory of goodness of fit," *Jour. Roy. Stat. Soc.*, Vol. 87 (1924), pp. 76–82.

[8] H. CRAMÉR, *Mathematical Methods of Statistics*, Princeton University Press, 1946, p. 424.

[9] J. NEYMAN, "Contribution to the theory of the $\chi^2$ test," *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1949, pp. 239–273.

[10] E. FIX, "Tables of the Noncentral $\chi^2$," *Univ. California Publ. Stat.*, Vol. 1 (1949), pp. 15–19.

[11] P. B. PATNAIK, "The non-central $\chi^2$- and $F$-distributions and their applications," *Biometrika*, Vol. 36 (1949), pp. 202–232.

[12] A. WALD, "Tests of statistical hypotheses concerning several parameters when the number of observations is large," *Trans. Am. Math. Soc.*, Vol. 54 (1943), pp. 426–482.

[13] G. A. BARNARD, "Significance tests for $2 \times 2$ tables," *Biometrika*, Vol. 34 (1947), pp. 123–138.

[14] R. A. FISHER, *The Design of Experiments*, Oliver and Boyd Ltd., Edinburgh, 1935.

[15] R. A. FISHER, *Statistical Methods for Research Workers*, Edinburgh, 5th and subsequent editions, Oliver and Boyd Ltd., Edinburgh, 1934, Section 21.02.

[16] K. D. TOCHER, "Extension of the Neyman-Pearson theory of tests to discontinuous variates," *Biometrika*, Vol. 37 (1950), pp. 130–144.

[17] D. LEWIS AND C. J. BURKE, "The use and misuse of the chi-square test," *Psych. Bull.*, Vol. 46 (1949), pp. 433–489.

[18] A. L. EDWARDS, "On the use and misuse of the chi-square test—the case of the $2 \times 2$ contingency table," *Psych. Bull.*, Vol. 47 (1950), pp. 341–346.

[19] W. G. COCHRAN, "The $\chi^2$ correction for continuity," *Iowa State Coll. Jour. Sci.*, Vol. 16 (1942), pp. 421–436.

[20] D. MAINLAND, "Statistical methods in medical research," *Canadian Jour. Res., E*, Vol. 26 (1948), pp. 1–166.

[21] R. A. FISHER AND F. YATES, *Statistical Tables for Biological, Agricultural and Medical Research*, Oliver and Boyd, Ltd., Edinburgh, 1938, Table VIII.

[22] J. NEYMAN AND E. S. PEARSON, "Further notes on the $\chi^2$ distribution," *Biometrika*, Vol. 22 (1931), pp. 298–305.

[23] W. G. COCHRAN, "The $\chi^2$ distribution for the binomial and Poisson series, with small expectations," *Annals of Eugenics*, Vol. 7 (1936), pp. 207–217.

[24] P. V. SUKHATME, "On the distribution of $\chi^2$ in small samples of the Poisson series," *Jour. Roy. Stat. Soc. Suppl.*, Vol. 5 (1938), pp. 75–79.

[25] G. H. FREEMAN AND J. H. HALTON, "Note on an exact treatment of contingency, goodness of fit and other problems of significance," *Biometrika*, Vol. 38 (1951), pp. 141–149.

[26] J. B. S. HALDANE, "The exact value of the moments of the distribution of $\chi^2$, used as a test of goodness of fit, when expectations are small," *Biometrika*, Vol. 29 (1937), pp. 133–143.

[27] J. B. S. HALDANE, "The mean and variance of $\chi^2$, when used as a test of homogeneity, when expectations are small," *Biometrika*, Vol. 31 (1939), pp. 346–355.

[28] F. YATES, "Contingency tables involving small numbers and the $\chi^2$ test," *Jour. Roy. Stat. Soc. Suppl.*, Vol. 1 (1934), pp. 217–235.

[29] H. B. MANN AND A. WALD, "On the choice of the number of class intervals in the application of the chi square test," *Annals of Math. Stat.*, Vol. 13 (1942), pp. 306–317.

[30] E. J. GUMBEL, "On the reliability of the classical $\chi^2$ test," *Annals of Math. Stat.*, Vol. 14 (1943), pp. 253–263.

[31] C. ARTHUR WILLIAMS, "On the choice of the number and width of classes for the chi-square test of goodness of fit," *Jour. Am. Stat. Assn.*, Vol. 45 (1950), pp. 77–86.

[32] J. BERKSON, "Some difficulties of interpretation encountered in the application of the chi-square test," *Jour. Am. Stat. Assn.*, Vol. 33 (1938), pp. 526–536.

[33] A. KOLMOGOROV, "Confidence limits for an unknown distribution function," *Annals of Math. Stat.*, Vol. 12 (1941), pp. 461–465.

[34] R. A. FISHER, "Has Mendel's work been re-discovered?" *Annals of Science*, Vol. 1 (1936), pp. 115–137.

[35] H. CRAMÉR, "On the composition of elementary errors," *Skandinavisk Aktuarietidskrift*, Vol. 11 (1928), pp. 13–74, 141–180.

[36] R. VON MISES, *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik*, Deuticke, Leipzig und Wien, 1931, pp. 316–335.

[37] N. SMIRNOV, "Sur la distribution de $\omega^2$," *C. R. Acad. Sci.* Paris, Vol. 202 (1936), p. 449.

[38] D. A. DARLING AND T. W. ANDERSON, "Asymptotic theory of certain goodness of fit criteria based on stochastic processes," *Annals of Math. Stat.*, Vol. 23 (1952), pp. 193–212.

[39] R. A. FISHER, "The significance of deviations from expectation in a Poisson series," *Biometrics*, Vol. 6 (1950), pp. 17–24.

[40] J. NEYMAN, "Smooth test for goodness of fit," *Skandinavisk Aktuarietidskrift*, Vol. 20 (1937), pp. 150–199.

[41] F. N. DAVID, "A $\chi^2$ 'smooth' test for goodness of fit," *Biometrika*, Vol. 34 (1947), pp. 299–310.

[42] H. O. LANCASTER, "The derivation and partition of $\chi^2$ in certain discrete distributions," *Biometrika*, Vol. 36 (1949), pp. 117–129.

[43] J. O. IRWIN, "A note on the subdivision of $\chi^2$ into components," *Biometrika*, Vol. 36 (1949), pp. 130–134.

[44] F. YATES, "The analysis of contingency tables with groupings based on quantitative characters," *Biometrika*, Vol. 35 (1948), pp. 176–181.

# SOME RANK ORDER TESTS WHICH ARE MOST POWERFUL AGAINST SPECIFIC PARAMETRIC ALTERNATIVES[1]

By Milton E. Terry

*University of North Carolina and Virginia Polytechnic Institute*

**Summary.** The most powerful rank order tests against specific parametric alternatives are derived.

Following the methods of Hoeffding [4], we derive the most powerful rank order test of whether $N$ observations come from the same but unknown population against the alternative that the observations $Z_1, \cdots, Z_N$ come from populations which have the joint density

$$\prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(z_i - d_i\xi - \eta)^2\right],$$

where $d_1, \cdots, d_N$ are given constants, not all equal, and $\xi/\sigma$ is sufficiently small. The test criterion was found to be $c_1(R) = \sum d_i E Z_{N,r_i}$, where $EZ_{Ni}$ is the expected value of the $i$th standard normal order statistic and $R = (r_1, \cdots, r_N)$ is the permutation of the ranks. The distribution of this statistic was shown to be asymptotically normal providing the known constants $d_1, \cdots, d_N$ satisfied Noether's condition [9].

The two-sample distribution is a special case, and the resultant statistic $c_1(R)$ is shown to be asymptotically normal. The approximation of the distribution of the $c_1(R)$ statistic to the distribution $C(1 - x^2)^{\frac{1}{2}N-2}$, $-1 \leqq x \leqq 1$, is investigated. This statistic is then compared to the existing Mann and Whitney $U$ statistic. No method having been found for analytical evaluation of the power of this test, the power was examined experimentally.

Tables are appended giving the exact distribution of the $c_1(R)$ statistic for all possible subsample sizes whose total size is less than or equal to 10 together with the corresponding Mann and Whitney $U$ value. Table 2 gives critical values of $c_1(R)$ for $N \leqq 10$, $p \leqq .10$.

**1. Introduction.** The problem of testing whether two samples come from the same population when either there are observed measurements and no assumption about the functional form of the underlying distributions or there are only ranks of the observations available has been treated by many statisticians.

H. Scheffé [12] gives an exhaustive and succinct review of this problem together with brief descriptions of proposed solutions offered previous to 1943. Included in Scheffé's references are the well known "runs" test and the modified likelihood ratio test of Wald and Wolfowitz [14], [17].

---

[1] This paper was presented at the Boston meeting of the Institute on December 26, 1951. These results were included in the thesis submitted to the Dept. of Mathematical Statistics of the University of North Carolina in partial fulfillment of the requirements for the degree of Doctor of Philosophy, June 1951.

A rank analogue to the standard parametric test was introduced in 1945 by Wilcoxon [16], and a linear function of Wilcoxon's $T$ was proposed by K. K. Matthen [7] in 1946 and its distribution and property of consistency shown by Mann and Whitney [6] in 1947. Tables are given by Mann and Whitney [6] for values of $m, n \leqq 8$.

Implicit in the introduction to Tables XX and XXI of Fisher and Yates [3] is a consideration by Fisher and Yates of rank order tests of the type discussed and developed in this dissertation. W. Hoeffding [4] develops a general most powerful rank order test, and this paper is an application of his results to specific problems, with various extensions.

**2. Derivation of the statistic $c_1(R)$.** Let $Z = (Z_1, \cdots, Z_N)$ be a random vector of $N$ components with probability function $P(S) = P\{Z \varepsilon S\}$, and $W$ be the set of points $(z_1, \cdots, z_N)$ in $N$-dimensional Euclidean space where $z_i \neq z_j$, $i \neq j$. We shall consider only probability functions such that

$$P(W) = 1. \tag{2.1}$$

Define $D_0$ to be a class of probability functions $P(S)$ which are invariant under all the $N!$ permutations of the $Z_i$ and $H_0$ to be the hypothesis that $P(S)$ is in $D_0$. In particular $H_0$ may be the hypothesis that $Z_1, Z_2, \cdots, Z_N$ are independent with a common continuous distribution.

Let $R = (r_1, \cdots, r_N)$ be a permutation of the integers $(1, \cdots, N)$. Let $S(R)$ be the subset of $W$ where $z_i$ has rank $r_i$, $(i = 1, \cdots, N)$. Let the integer $t_i$ be defined by $r_{t_i} = i$, $(i = 1, \cdots, N)$. Then $S(R)$ is the set where $z_{t_1} < z_{t_2} < \cdots < z_{t_N}$. The set $W$ consists of the $M = N!$ subsets $S(R)$. For the probability $P(S(R))$ of the set $S(R)$ we shall use the short notation $P(R)$. A probability function $P(S)$ which is in $D_0$ will be denoted by $P(S \mid H_0)$. Let $H_1$ be the hypothesis that $P(A) = P_1(A)$, a probability function of $Z$ not in $D_0$. Denote the $M$ permutations $R$ by $R_1, \cdots, R_M$ in such a way that

$$P_1(R_i) \geqq P_1(R_j) \qquad \text{if } i = 1, \cdots, m; j = m + 1, \cdots, M.$$

Clearly, $(R_1, \cdots, R_m)$ determines a rank order test which is most powerful for testing $H_0$ against $H_1$ [4].

If the condition (2.1) is not satisfied, we may apply either of the following two rules.

*Rule* 1. If $n$ of the values $Z_1, \cdots, Z_N$ are less than $Z_j$, and $m$ values are equal to $Z_j$, the latter values are assigned at random the ranks $n + 1, \cdots, n + m$. This is carried out for all $Z_j$, and $Z$ is treated as belonging to the corresponding set $S(R)$.

*Rule* 2. Let $k$ be the number of sets $S(R)$ which can be obtained by applying 1 to a given $Z$. If exactly $k'$ of these belong to the critical region, $H_0$ is rejected with probability $k'/k$.

It follows from the definition of $D_0$ that

$$P(R \mid H_0) = \frac{1}{N!} \tag{2.2}$$

for all $R$.

Consider the alternative where $Z_1, \cdots, Z_N$ have the joint density

$$\prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(z_i - d_i\xi - \eta)^2\right], \tag{2.3}$$

where $d_1, \cdots, d_N$ are given constants, not all equal, and $\xi$, $\eta$ are parameters. Then

$$P(R) = \int_{z_{t_1} < \cdots < z_{t_N}} \cdots \int \prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(z_i - d_i\xi - \eta)^2\right] dz_1 \cdots dz_N. \tag{2.4}$$

The probability $P(R)$ is independent of $\eta$ and depends on $\delta = \xi/\sigma$ only. For if we make the transformation

$$z_i' = \frac{1}{\sigma}(z_i - \eta), \qquad (i = 1, \cdots, N),$$

the inequalities $z_{t_1} < z_{t_2} < \cdots < z_{t_N}$ are transformed into $z'_{t_1} < z'_{t_2} < \cdots < z'_{t_N}$. Denoting $P(R)$ by $P(R, \delta)$, and replacing $z_i'$ by $z_i$, we have

$$P(R, \delta) = \int_{z_{t_1} < \cdots < z_{t_N}} \cdots \int \prod_{i=1}^{N} f(z_i - d_i\delta)\, dz_1 \cdots dz_N, \tag{2.5}$$

where

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right). \tag{2.6}$$

It is easily seen that $P(R, \delta)$ has continuous derivatives of any order with respect to $\delta$. Hence we can write, for any positive integer $k$,

$$P(R, \delta) = \frac{1}{N!}\left[c_0(R) + c_1(R)\frac{\delta}{1} + \cdots + c_k(R)\frac{\delta^k}{k!} + O(\delta^{k+1})\right], \tag{2.7}$$

where

$$\frac{1}{N!}c_\eta(R) = \frac{d^\eta(P(R, \delta))}{d\delta^\eta}\bigg|_{\delta=0},$$

and $O(\delta^{k+1})$ denotes a term of order $\delta^{k+1}$ as $\delta \to 0$. We shall also write formally

$$P(R, \delta) = \frac{1}{N!}\sum_{\eta=0}^{\infty} c_\eta(R)\frac{\delta^\eta}{\eta!} \tag{2.8}$$

without considering the convergence of the series.

To obtain the coefficients $c_\eta(R)$ we shall expand the integrand in (2.5) in powers of $\delta$ and integrate term by term.

$$f(z_i - d_i\delta) = \sum_{\eta_i=0}^{\infty} \frac{f^{(\eta_i)}(z_i)(-d_i\delta)^{\eta_i}}{\eta_i!}, \tag{2.9}$$

where $f^{(\eta_i)}(z_i)$ is the $\eta_i$th derivative of $f(z_i)$. Then

$$\prod_{i=1}^{N} f(z_i - d_i\delta) = \prod_{i=1}^{N} \sum_{\eta_i=0}^{\infty} \frac{f^{(\eta_i)}(z_i)(-d_i\delta)^{\eta_i}}{\eta_i!}$$

$$= \sum_{\eta_1=0}^{\infty} \cdots \sum_{\eta_N=0}^{\infty} \frac{(-\delta)^{\exp\left[\sum_{i=1}^{N} \eta_i\right]}}{\eta_1!\,\eta_2! \cdots \eta_N!} \prod_{i=1}^{N} f^{(\eta_i)}(z_i)\, d_i^{\eta_i} \tag{2.10}$$

$$= \sum_{\eta=0}^{\infty} \sum^{(\eta)} \frac{\eta!}{\prod_{i=1}^{N} \eta_i!} \prod_{i=1}^{N} d_i^{\eta_i} f^{(\eta_i)}(z_i)\, \frac{(-\delta)^{\eta}}{\eta!},$$

where $\sum^{(\eta)}$ represents the summation over all $\eta_1, \cdots, \eta_N$ such that $\eta_i \geqq 0$, $i = 1, \cdots, N$, and $\sum_{i=1}^{N} \eta_i = \eta$. Integrating the final form of (2.10) term by term and comparing the resultant expansion with (2.8) we obtain

$$c_\eta(R) = N! \int \cdots \int_{z_{i_1} < \cdots < z_{i_N}} \sum^{(\eta)} \frac{\eta!}{\prod_{i=1}^{N} \eta_i!} \prod_{i=1}^{N} d_i^{\eta_i} (-1)^{\eta_i} f^{(\eta_i)}(z_i)\, dz_1 \cdots dz_N. \tag{2.11}$$

Consider the following transformation. Let $z_{i_i} = z_i'$, $i = 1, \cdots, N$, and now the domain of integration will be $(z_1' < \cdots < z_N')$. Since by definition $t_{r_i} = i$, $i = 1, \cdots, N$, where $r_i$ is the rank of $z_i$ in $S(R)$, we have $z_i = z_{t_{r_i}} = z_{r_i}'$. Applying this transformation to $c_\eta(R)$ in (2.11) and replacing $z_i'$ by $z_i$, we have

$$c_\eta(R) = N! \int \cdots \int_{z_1 < \cdots < z_N} \sum^{(\eta)} \frac{\eta!}{\prod_{i=1}^{N} \eta_i!} \prod_{i=1}^{N} (-d_i)^{\eta_i} f^{(\eta_i)}(z_{r_i})\, dz_1 \cdots dz_N. \tag{2.12}$$

But

$$f^{(\eta)}(z) = (-1)^{\eta} H_\eta(z) f(z), \tag{2.13}$$

where $H_\eta(z)$ is the $\eta$th order Hermite polynomial [1]. In particular, $H_0(z) = 1$, $H_1(z) = z$, $H_2(z) = z^2 - 1$. Substituting the appropriate forms of (2.13) in (2.12) we have

$$c_\eta(R) = N! \int \cdots \int_{z_1 < \cdots < z_N} \sum^{(\eta)} \frac{\eta!}{\prod_{i=1}^{N} \eta_i!} \prod_{i=1}^{N} d_i^{\eta_i} H_{\eta_i}(z_{r_i}) f(z_{r_i})\, dz_1 \cdots dz_N. \tag{2.14}$$

Let $Z_{N1} \leqq Z_{N2} \leqq \cdots \leqq Z_{NN}$ be the order statistics in a random sample of $N$ from a standard normal population. The joint frequency function of $Z_{N1}, \cdots, Z_{NN}$ is $N! \prod_{i=1}^{N} f(z_i)$, $z_1 < \cdots < z_N$. Hence

$$c_\eta(R) = E\left[ \sum^{(\eta)} \frac{\eta!}{\prod_{i=1}^{N} \eta_i!} \prod_{i=1}^{N} d_i^{\eta_i} H_{\eta_i}(Z_{Nr_i}) \right]. \tag{2.15}$$

In particular,

$$c_0(R) = 1, \tag{2.16}$$

$$c_1(R) = \sum_{i=1}^{N} d_i E(Z_{Nr_i}), \tag{2.17}$$

$$\begin{aligned} c_2(R) &= E\left[\sum_{i=1}^{N} d_i Z_{Nr_i}\right]^2 - \sum_{i=1}^{N} d_i^2, \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} d_i d_j E(Z_{Nr_i} Z_{Nr_j}) - \sum_{i=1}^{N} d_i^2. \end{aligned} \tag{2.18}$$

To the order of $\delta^3$, $P(R, \delta)$ may be written as follows:

$$P(R, \delta) = \frac{1}{N!}\left[1 + c_1(R)\delta + c_2(R)\frac{\delta^2}{2} + O(\delta^3)\right]. \tag{2.19}$$

If $\delta$ is positive and sufficiently small, the rank order test of size $K/N!$ which is most powerful against $P(S, \delta)$ is determined by the $K$ permutations $R_i$, $i = 1, \cdots, K$, for which $c_1(R)$ takes its largest values. Hence $c_1(R)$ is the desired statistic for the stated alternative with $\delta$ positive and small.

**3. Asymptotic distribution.** It will be proved that under $H_0$ the statistic $c_1(R)$ is asymptotically normally distributed as $N \to \infty$ if the $d_i$ satisfy a certain condition. The proof is based on a theorem of Noether [9] and an unpublished theorem of Hoeffding which, together with its proof, is reproduced here by permission.

A sequence of sequences $A_N = (a_{N1}, \cdots, a_{NN})$, $N = 1, 2 \cdots$ is said to satisfy Condition $W$ if as $N \to \infty$,

$$\frac{\frac{1}{N}\sum_{i=1}^{N}(a_{Ni} - \bar{a}_N)^r}{\left[\frac{1}{N}\sum_{i=1}^{N}(a_{Ni} - \bar{a}_N)^2\right]^{r/2}} = O(1), \qquad r = 3, 4, \cdots, \tag{3.1}$$

where $\bar{a}_N = 1/N \sum_{i=1}^{N} a_{Ni}$ [15]. Similarly a sequence of sequences satisfies Noether's condition [9] if

$$\frac{\sum(a_{Ni} - \bar{a}_N)^r}{[\sum(a_{Ni} - \bar{a}_N)^2]^{r/2}} = o(1), \qquad r = 3, 4, \cdots. \tag{3.2}$$

Let $D_N = (d_{N1}, \cdots, d_{NN})$ and $H_N = (h_{N1}, \cdots, h_{NN})$, $N = 1, 2, \cdots$, be two sequences of sequences of real numbers. Let $(x_1, \cdots, x_N)$ be the random vector whose domain consists of the $N!$ equally probable permutations of $(h_{N1}, \cdots, h_{NN})$, and let

$$L_N = d_{N1}x_1 + \cdots + d_{NN}x_N,$$

$$L_N^0 = \frac{L_N - E_0(L_N)}{\sqrt{V_0(L_N)}},$$

where $E_0(L_N)$ and $V_0(L_N)$ denote the mean and the variance of $L_N$.

Noether [9] shows that

$$E_0(L_N) = \frac{1}{N}\sum_{i=1}^{N} d_{Ni}\sum_{i=1}^{N} h_{Ni} = N\bar{d}_N\bar{h}_N, \tag{3.3}$$

$$V_0(L_N) = \frac{1}{N-1}\sum_{i=1}^{N}(d_{Ni}-\bar{d}_N)^2\sum_{i=1}^{N}(h_{Ni}-\bar{h}_N)^2, \tag{3.4}$$

and that if the sequence $H_N$ satisfies Condition $W$, and the $D_N$ satisfies Noether's condition, $L_N^0$ has a normal limiting distribution with mean 0 and variance 1.

The statistic $c_1(R)$ (2.17) can be written

$$c_1(R) = \sum_{i=1}^{N} d_i h_{r_i},$$

where

$$h_i = h_{Ni} = E(Z_{Ni}). \tag{3.5}$$

(The $d_i = d_{Ni}$ may also depend on $N$; to simplify the notation we write $d_i$, $h_i$ for $d_{Ni}$, $h_{Ni}$). Under $H_0$ the $N!$ permutation of $(h_1, \cdots, h_N)$ are equally probable, so that $c_1(R) = L_N$, with $h_{Ni}$ defined by (3.5).

Hoeffding shows that the sequences $H_N = (h_1, \cdots, h_N) = (EZ_{N1}, \cdots, EZ_{NN})$ satisfy Condition $W$ as follows.

It is easily seen from the symmetry of the normal distribution that $h_i = EZ_{Ni} = -EZ_{N,N+1-i} = -h_{n+1-i}$. Hence

$$\sum_{i=1}^{N} h_i^r = 0 \text{ if } r \text{ is odd.} \tag{3.6}$$

In particular $\bar{h}_N = 0$. Hence it suffices to show that

$$\frac{\frac{1}{N}\sum_{i=1}^{N} h_i^{2k}}{\left[\frac{1}{N}\sum_{i=1}^{N} h_i^2\right]^k} = O(1) \qquad \text{for } k = 2, 3, \cdots.$$

If $r = 2k$ is even, we have

$$\sum_{i=1}^{N} h_i^{2k} = \sum_{i=1}^{N}(EZ_{Ni})^{2k} \leqq \sum_{i=1}^{N} E(Z_{Ni})^{2k} = E\sum_{i=1}^{N} Z_{Ni}^{2k}, \tag{3.7}$$

and

$$E\sum Z_{Ni}^{2k} = NEX^{2k}, \tag{3.8}$$

where $X$ is standard normal. Equation (3.8) holds since a symmetrical function of the order statistics is distributed as the same function of the unordered independent variables.

If $m = [N/2]$ is the largest integer $\leqq N/2$, we have by Cauchy's inequality

$$\sum_{i=1}^{N} h_i^2 = 2 \sum_{i=N-m+1}^{N} h_i^2 \geqq \frac{2}{m}\left(\sum_{i=N-m+1}^{N} h_i\right)^2 = \frac{2}{m}\left(\sum_{i=N-m+1}^{N} EZ_{Ni}\right)^2. \tag{3.9}$$

Let $F(x) = \int_{-\infty}^{x} f(y)\,dy$. Then for $1 \leqq k \leqq N-1$,

$$\begin{aligned}\sum_{i=k+1}^{N} EZ_{Ni} &= \sum_{i=k+1}^{N} N\binom{N-1}{i-1}\int_{-\infty}^{\infty} xf(x)[F(x)]^{i-1}[(1-F(x))]^{N-1}\,dx \\ &= N\int_{-\infty}^{\infty} xf(x)\sum_{j=k}^{N-1}\binom{N-1}{j}[F(x)]^j[1-F(x)]^{N-1-j}\,dx.\end{aligned} \tag{3.10}$$

Since $\sum_{j=k}^{N-1}\binom{N-1}{j}[F(x)]^j[1-F(x)]^{N-1-j}$ is the remainder after $k$ terms of the binomial $(F(x)+1-F(x))^{N-1}$, we may express [5] this remainder as the incomplete Beta function

$$\frac{(N-1)!}{(k-1)!(N-1-k)!}\int_0^{F(x)} t^{k-1}(1-t)^{N-1-k}\,dt.$$

Hence

$$\begin{aligned}&\sum_{i=k+1}^{N} EZ_{Ni} \\ &\quad = N\int_{-\infty}^{\infty} xf(x)\,\frac{(N-1)!}{(k-1)!(N-1-k)!}\int_0^{F(x)} t^{k-1}(1-t)^{N-1-k}\,dt\,dx.\end{aligned} \tag{3.11}$$

Integrating by parts we obtain

$$\begin{aligned}&\sum_{i=k+1}^{N} EZ_{Ni} \\ &\quad = \frac{N!}{(k-1)!(N-k-1)!}\int_{-\infty}^{\infty}[f(x)]^2[(Fx)]^{k-1}[1-F(x)]^{N-1-k}\,dx.\end{aligned} \tag{3.12}$$

Since $F(-x) = 1-F(x) \leqq x^{-1}f(x)$ for $x > 0$, there exists a constant $K > 0$ such that

$$f(x) \geqq KF(x)[1-F(x)] \text{ for all } x. \tag{3.13}$$

For let $c > 0$; then

$$d = \min_{x<c}\frac{f(x)}{F(x)[1-F(x)]} > 0.$$

For $x > c$, $f(x) > x[1-F(x)] \geqq x[1-F(x)][F(x)] \geqq cF(x)[1-F(x)]$; and by symmetry the same inequality holds for $x < c$. Hence we can take $K = \min(c, d)$.

$$\begin{aligned}\sum_{i=k+1}^{N} EZ_{Ni} &\geqq K\,\frac{N!}{(k-1)!(N-1-k)!}\int_{-\infty}^{\infty} f(x)[F(x)]^k[1-F(x)]^{N-k}\,dx \\ &= K\,\frac{k(N-k)}{N+1}.\end{aligned} \tag{3.14}$$

Whence from (3.9)

$$\sum_{i=1}^{N} h_i^2 \geqq \frac{2}{m}\left(K\frac{(N-m)m}{N+1}\right)^2 \geqq k'N, \tag{3.15}$$

where $k'$ is a positive constant. Then by (3.7), (3.8) and (3.15),

$$\frac{1}{N}\sum_{i=1}^{N} h_i^2\left(\frac{1}{N}\sum_{i=1}^{N} h_i^2\right)^{-k} \leqq Ex^{2k}(k')^{-k} = O(1), \tag{3.16}$$

whence it follows that $EZ_{N1}, \cdots, EZ_{NN}$ satisfies Condition $W$. By (3.3), (3.4), and (3.5) we have

$$E_0(c_1(R)) = 0,$$

$$V_0(c_1(R)) = E_0(c_1(R))^2 = \frac{1}{N-1}\sum_{i=1}^{N}(d_i - \bar{d})^2 \sum_{i=1}^{N} h_i^2. \tag{3.17}$$

We now obtain from the theorem of Noether:

THEOREM. *If the sequences* $(d_1, \cdots, d_N)$, $N = 1, 2, \cdots$ *satisfy Neother's condition, then under* $H_0$

$$\frac{c_1(R)}{\sqrt{E_0(c_1(R))^2}} = \frac{\sum_{i=1}^{N} d_i EZ_{Nr_i}}{\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(d_i - \bar{d})^2 \sum_{i=1}^{N}(EZ_{Ni})^2}}$$

*has a limiting normal distribution with mean* 0 *and variance* 1.

**4. Approximate distribution.** Pitman [10], [11] has proposed the following approximation of the distribution of a statistic of the type of $c_1(R)$ when the permutations are equally probable. We now assume without loss of generality that $\sum_{i=1}^{N} = d_i = 0$. Let

$$r = r(R) = \frac{\sum_{i=1}^{N} d_i h_{r_i}}{\sqrt{\sum_{i=1}^{N} d_i^2 \sum_{i=1}^{N} h_i^2}} = \frac{c_1(R)}{\sqrt{\sum_{i=1}^{N} d_i^2 \sum_{i=1}^{N}(EZ_{Ni})^2}}.$$

We shall first show that the odd moments of $c_1(R)$, and hence those of $r$, are zero. If $R = (r_1, \cdots, r_N)$, let $R' = (N+1-r_1, \cdots, N+1-r_N)$. Let $h_i = EZ_{Ni}$, then $-h_i = h_{N+1-i}$; and we have

$$c_1(R) = \sum_{i=1}^{N} d_i h_{r_i} = -\sum_{i=1}^{N} d_i h_{N+1-r_i} = -c_1(R') \quad \text{for all } R.$$

From the above it follows that

$$c_1(R)^{2k+1} = -c_1(R')^{2k+1}. \tag{4.1}$$

Hence

$$E_0[c_1(R)]^{2k+1} = \frac{1}{N!}\sum_{R} c_1(R)^{2k+1} = -\frac{1}{N!}\sum_{R} c_1(R')^{2k+1}. \tag{4.2}$$

Since summation over all $R$ is equivalent to summation over all $R'$, we obtain

$$E_0[c_1(R)]^{2k+1} = -\frac{1}{N!}\sum_{R'} c_1(R')^{2k+1} = -E_0[c_1(R)]^{2k+1}, \tag{4.3}$$

and thus

$$E_0[c_1(R)]^{2k+1} = 0, \qquad E_0(r^{2k+1}) = 0. \tag{4.4}$$

It follows from (3.16) and from Pitman's results that

$$E_0(r^2) = \frac{1}{N-1}, \tag{4.5}$$

$$E_0(r^4) = \frac{3}{(N-1)(N+1)}\left[1 + \frac{(N-2)(N-3)}{3N(N-1)^2}\cdot\frac{k_4}{k_2^2}\cdot\frac{k_4'}{k_2'^2}\right], \tag{4.6}$$

where

$$k_2 = \frac{1}{N-1}\sum_{i=1}^{N} d_i^2$$

$$k_4 = \frac{N}{(N-1)(N-2)(N-3)}\left[(N+1)\sum_{i=1}^{N} d_i^4 - 3\frac{N-1}{N}\left(\sum_{i=1}^{N} d_i^2\right)^2\right],$$

and $k_2'$, $k_4'$ are the corresponding functions of the $h_i$ ($k_i$ and $k_i'$ are Fisher's $k$-statistics).

Since $r$ has the form of a correlation coefficient, we have $-1 \leqq r \leqq 1$, and $E_0(r^{2k+1}) = 0$, $E_0(r^2) = 1/(N-1)$. If the last term in the square brackets of (4.6) is small compared with 1, we also have $E_0(r^4) = 3/[(N-1)(N+1)]$ approximately.

The distribution with the frequency function

$$g(x) = \begin{cases} \dfrac{1}{B(\frac{1}{2}, \frac{1}{2}N-1)}(1-x^2)^{\frac{1}{2}N-2}, & -1 \leqq x \leqq 1 \\ 0, & \text{elsewhere} \end{cases} \tag{4.7}$$

has its odd moments equal to zero, and its second and fourth moments, respectively $1/(N-1)$ and $3/[(N-1)(N+1)]$.

For this reason the distribution $g(x)$ may be a suitable approximation of the distribution of $r$ under $H_0$. In Section 7 it is shown that, in the case where the $d_i$ take on two values only (two-sample test), the approximation is satisfactory even for small values of $N$. Since the distribution $g(x)$ may be a suitable approximation to the distribution of $r$, we can expect the Student's distribution with $N-2$ degrees of freedom also to approximate the distribution of $(r/\sqrt{1-r^2})\sqrt{N-2}$. Thus we can obtain the approximate critical values of $c_1(R)$ from the $t$ tables.

**5. Two-sample statistic $c_1(R)$.** Let $H_0$ be the hypothesis that two samples of $m$ and $n$ observations come from the same continuous population. In this sec-

tion the rank order test of $H_0$ will be considered which is most powerful against the alternative that the two samples come from two normal populations with common variance $\sigma^2$ and different means $\mu_{1.1}$ and $\mu_{1.2}$ with $\delta = (\mu_{1.2} - \mu_{1.1})/\sigma$ sufficiently small.

The hypothesis $H_0$ implies that the joint distribution of the $N = m + n$ variables is invariant under all permutations, and that $P(W)$, the probability that all $N$ values are different, is one. Hence the assumptions about $H_0$ in Section 2 are satisfied. The alternative is a special case of the distribution (2.3) with

$$d_i\xi + \eta = \mu_{1.1}, \qquad i = 1, \cdots, m,$$

$$d_{m+j}\xi + \eta = \mu_{1.2}, \qquad j = 1, \cdots, n.$$

If $d_i = -n/(n + m)$, $i = 1, \cdots, m$; $d_{m+j} = m/(n + m)$, $j = 1, \cdots, n$, then $d = \sum_{i=1}^{N} d_i/N = 0$,

$$\xi = \mu_{1.2} - \mu_{1.1},$$

$$\eta = \frac{m\mu_{1.1} + n\mu_{1.2}}{m + n}, \tag{5.1}$$

$$\delta = \frac{\xi}{\sigma} = \frac{\mu_{1.2} - \mu_{1.1}}{\sigma}.$$

Then, from (2.17),

$$\begin{aligned} c_1(R) &= -\frac{n}{N}\sum_{i=1}^{m} E(Z_{Nr_i}) + \frac{m}{N}\sum_{i=m+1}^{N} E(Z_{Nr_i}) \\ &= -\frac{n}{N}\sum_{i=1}^{N} E(Z_{Nr_i}) + \sum_{i=m+1}^{N} E(Z_{Nr_i}) \qquad (5.3) \\ &= \sum_{j=1}^{n} E(Z_{Nr_{m+j}}). \end{aligned}$$

**6. Asymptotic distribution of two-sample $c_1(R)$.**

THEOREM. *If $H_0$ is true, the variance of $c_1(R)$ is*

$$V_0(c_1(R)) = \frac{mn}{N(N-1)}\sum_{i=1}^{N} (EZ_{Ni})^2. \tag{6.1}$$

*If there exist two constants $k$, $k'$, $0 < k < k' < 1$, such that $k < n/N < k'$ as $N \to \infty$, then*

$$\frac{c_1(R)}{\sqrt{\dfrac{mn}{N(N-1)}\displaystyle\sum_{i=1}^{N} (EZ_{Ni})^2}}$$

*has a normal limiting distribution with mean 0 and variance 1.*

PROOF. Relation (6.1) follows from (3.16) since

$$\sum_{i=1}^{N} d_{Ni} - \bar{d}_N)^2 = \sum_{i=1}^{N} d_{Ni}^2 = \frac{mn^2}{(m+n)^2} + \frac{nm^2}{(m+n)^2} = \frac{mn}{m+n} = \frac{mn}{N}.$$

By the theorem of Section 2 it now suffices to show that

$$\frac{\sum_{i=1}^{N} d_{Ni}^r}{\left[\sum_{i=1}^{N} d_{Ni}^2\right]^{r/2}} = o(1), \qquad r = 3, 4, \cdots.$$

This follows from the relations

$$\sum_{i=1}^{N} d_{Ni}^2 = \frac{mn}{N} \tag{6.2}$$

and

$$\sum_{i=1}^{N} d_{Ni}^r = m\left(-\frac{n}{N}\right)^r + n\left(\frac{m}{N}\right)^r = N^{-r}[nm^r + (-1)^r mn^r]. \tag{6.3}$$

Hence

$$\frac{\sum_{i=1}^{N} d_{Ni}^r}{\left[\sum_{i=1}^{N} d_{Ni}^2\right]^{r/2}} = n\left(\frac{m}{nN}\right)^{r/2} + (-1)^r m\left(\frac{n}{mN}\right)^{r/2} = o(1), \qquad r = 3, 4, \cdots,$$

since the inequality $0 < k < n/N < k' < 1$ implies $0 < 1 - k' < m/N < 1 - k < 1$.

**7. Approximate distribution of two-sample $c_1(R)$.** In Section 4 we have considered the approximation of the distribution under $H_0$ of

$$r = \frac{c_1(R)}{\sqrt{\sum_{i=1}^{N} d_i^2 \sum_{i=1}^{N} (EZ_{Ni})^2}}$$

by the distribution (4.7). In the two-sample case

$$r = \frac{c_1(R)}{\sqrt{\frac{mn}{N}\sum_{i=1}^{N} (EZ_{Ni})^2}},$$

$$E_0(r^2) = \frac{1}{N-1}, \tag{7.1}$$

and

$$E_0(r^4) = \frac{3}{N^2-1}\left[1 + \frac{(N+1)^2}{3N(N-2)(N-3)}\left(\frac{(m-n)^2}{mn} + \frac{6}{N+1}\right)\left(\frac{\mu_4}{\mu_2^2} - \frac{3(N-1)}{N+1}\right)\right].$$

Thus when the second term within the square brackets is small compared with 1, the distribution (4.7) can be expected to approximate the distribution of $r$.

If $X$ has the distribution $g(x)$, (4.7), then $X\sqrt{N-2}/\sqrt{1-X^2}$ has Student's $t$-distribution with $N$-2 degrees of freedom. Hence the latter distribution will approximate the distribution of $r\sqrt{N-2}/\sqrt{1-r^2}$. The $t$ approximation to the critical values of $c_1(R)$ at the level $\alpha$ of the $c_1(R)$ test for given $N$, $n$, $m$, $\mu_2$ is

$$c_{1,\alpha} = \sqrt{\frac{t^2_{N-2,2\alpha}\, mn\mu_2}{t^2_{N-2,2\alpha} + m + n - 2}} \tag{7.2}$$

where $t_{N-2,2\alpha}$ is the tabular value of $t$ with $N-2$ degrees of freedom at the level $2\alpha$ in the Fisher and Yates Table III [3].

The $t$ approximation is very good for $N$ as small as 6 insofar as comparisons of fourth moments and critical values are concerned. Shown below are the critical values, $c_{1,\alpha}$, as obtained from the exact distributions for the level of significance $\alpha \cong .05$ and from the $t$ approximation for the level $\alpha = .05$, together with $P\{c\ (R) \geqq c_{1,\alpha}\}$.

*A comparison of critical values*

| $N$ | $m$ | $n$ | $c_{1,\alpha}$ | | $P\{c_1(R) \geqq c_{1,\alpha}\}$ | |
|---|---|---|---|---|---|---|
| | | | Exact | $t$ | Exact | $t$ |
| 6 | 3 | 3 | 1.81 | 2.11 | .05 | .05 |
| 7 | 2 | 5 | 1.80 | 2.11 | .05 | .05 |
| 7 | 3 | 4 | 1.97 | 2.11 | .057 | .057 |
| 8 | 3 | 5 | 2.12 | 2.08 | .053 | .053 |
| 8 | 4 | 4 | 2.27 | 2.15 | .057 | .057 |
| 9 | 3 | 6 | 2.33 | 2.17 | .047 | .047 |
| 9 | 4 | 5 | 2.42 | 2.29 | .048 | .056 |
| 10 | 4 | 6 | 2.54 | 2.40 | .048 | .052 |
| 10 | 5 | 5 | 2.58 | 2.45 | .048 | .048 |

**8. Exact distribution of two-sample $c_1(R)$ for small $N$.** The exact distribution of $c_1(R)$ has been computed for small values of $N$, $n$, $m$; $N \leqq 10$ using the Fisher and Yates tables [3]. These distributions are displayed in Table 1, and the values of $c_1(R)$ and $U$ (cf. Sec. 9) are shown for the ${}_NC_n$ distinct permutations of $(x_1, \cdots, x_m, y_1, \cdots, y_n)$, $m + n = N$. Since the probability of each such permutation is $1/{}_NC_n$, these probabilities are not explicitly stated in the table. Since only the relation between $x$ and $y$ matters, we will replace each $x_i$ by a 0 and each $y_j$ by a 1, following the simplifying notation of Wolfowitz [14]. Each $R$ will thus be represented by a sequence of $m$ 0's and $n$ 1's. As the values of $c_1(R)$ for $N$, $m$, $n$ and $N$, $n$, $m$ are symmetric, the table will show the distribution for subsample sizes $m$, $n$; $m \leqq n$. The probability under the null hypothesis that a sample of $m$ 0's and $n$ 1's will give rise to a value of $c_1(R)$ exceeding or equaling the tabulated value of $c_1(R)$ is given in Table 2 for values of $p \leqq 0.10$, and is available for tests of significance when $6 \leqq N \leqq 10$.

TABLE 1

*Exact distribution of $C_1(R)$*

Exact distribution of $c_1(R)$ for distinct permutations of $R$, for given $N$, $m$, $n$; $m \leqq n$, together with corresponding values of Mann-Whitney $U$ test [6]. The value of $c_1(R)$ is tabulated under the column heading $c_1$.

| $R$ | $c_1$ | $U$ | $R$ | $c_1$ | $U$ | $R$ | $c_1$ | $U$ | $R$ | $c_1$ | $U$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $N = 2$ | $m = 1$ | | 001110 | .64 | 3 | 0110110 | .00 | 6 | 11000111 | .47 | 6 |
| 01 | .56 | 0 | 011001 | .43 | 4 | $N = 8$ | $m = 2$ | | 10011101 | .47 | 6 |
| $N = 3$ | $m = 1$ | | 010110 | .20 | 4 | 00111111 | 2.27 | 0 | 01110101 | .42 | 7 |
| 011 | .85 | 0 | 100101 | .20 | 4 | 01011111 | 1.89 | 1 | 10110011 | .23 | 7 |
| 101 | .00 | 1 | $N = 7$ | $m = 1$ | | 01101111 | 1.57 | 2 | 01101110 | .15 | 7 |
| $N = 4$ | $m = 1$ | | 0111111 | 1.35 | 0 | 10011111 | 1.32 | 2 | | | |
| 0111 | 1.03 | 0 | 1011111 | .76 | 1 | 01110111 | 1.27 | 3 | 11001011 | .15 | 7 |
| 1011 | .30 | 1 | 1101111 | .35 | 2 | | | | 10101101 | .15 | 7 |
| $N = 4$ | $m = 2$ | | 1110111 | .00 | 3 | 10101111 | 1.00 | 3 | 01111001 | .10 | 8 |
| 0011 | 1.33 | 0 | $N = 7$ | $m = 2$ | | 01111011 | .95 | 4 | $N = 8$ | $m = 4$ | |
| 0101 | .73 | 1 | 0011111 | 2.11 | 0 | 10110111 | .70 | 4 | 00001111 | 2.89 | 0 |
| 0110 | .00 | 2 | 0101111 | 1.70 | 1 | 11001111 | .62 | 4 | 00010111 | 2.59 | 1 |
| $N = 5$ | $m = 1$ | | 0110111 | 1.35 | 2 | 01111101 | .57 | 5 | 00100111 | 2.27 | 2 |
| 01111 | 1.66 | 0 | 1001111 | 1.11 | 2 | | | | 00011011 | 2.27 | 2 |
| 10111 | .50 | 1 | 0111011 | 1.00 | 3 | 10111011 | .38 | 5 | 00101011 | 1.95 | 3 |
| 11011 | .00 | 2 | | | | 11010111 | .32 | 5 | | | |
| $N = 5$ | $m = 2$ | | 1010111 | .76 | 3 | 11100111 | .00 | 6 | 01000111 | 1.89 | 3 |
| 00111 | 1.66 | 0 | 0111101 | .59 | 4 | 11011011 | .00 | 6 | 00011101 | 1.89 | 3 |
| 01011 | 1.16 | 1 | 1011011 | .41 | 4 | 10111101 | .00 | 6 | 00110011 | 1.65 | 4 |
| 01101 | .66 | 2 | 1100111 | .35 | 4 | 01111110 | .00 | 6 | 01001011 | 1.57 | 4 |
| 10011 | .50 | 2 | 1101011 | .00 | 5 | $N = 8$ | $m = 3$ | | 00101101 | 1.57 | 4 |
| 01110 | .00 | 3 | 1011101 | .00 | 5 | 00011111 | 2.74 | 0 | | | |
| 10101 | .00 | 3 | $N = 7$ | $m = 3$ | | 00101111 | 2.42 | 1 | 10000111 | 1.32 | 4 |
| $N = 6$ | $m = 1$ | | 0001111 | 2.46 | 0 | 00110111 | 2.12 | 2 | 00011110 | 1.32 | 4 |
| 011111 | 1.27 | 0 | 0010111 | 2.11 | 1 | 01001111 | 2.04 | 2 | 01010011 | 1.27 | 5 |
| 101111 | .64 | 1 | 0011011 | 1.76 | 2 | 00111011 | 1.80 | 3 | 00110101 | 1.27 | 5 |
| 110111 | .20 | 2 | 0100111 | 1.70 | 2 | | | | 01001101 | 1.19 | 5 |
| $N = 6$ | $m = 2$ | | 0011101 | 1.35 | 3 | 01010111 | 1.74 | 3 | | | |
| 001111 | 1.91 | 0 | | | | 10001111 | 1.47 | 3 | 10001011 | 1.00 | 5 |
| 010111 | 1.47 | 1 | 0101011 | 1.35 | 3 | 00111101 | 1.42 | 4 | 00101110 | 1.00 | 5 |
| 011011 | 1.07 | 2 | 1000111 | 1.11 | 3 | 01100111 | 1.42 | 4 | 01100011 | .95 | 6 |
| 100111 | .84 | 2 | 0110011 | 1.00 | 4 | 01011011 | 1.42 | 4 | 00111001 | .95 | 6 |
| 011101 | .63 | 3 | 0101101 | .94 | 4 | | | | 01010101 | .89 | 6 |
| | | | 1001011 | .76 | 4 | 10010111 | 1.17 | 4 | | | |
| 101011 | .44 | 3 | | | | 01101011 | 1.10 | 5 | 10010011 | .70 | 6 |
| 011110 | .00 | 4 | 0011110 | .76 | 4 | 01011101 | 1.04 | 5 | 00110110 | .70 | 6 |
| 101101 | .00 | 4 | 0110101 | .59 | 5 | 00111110 | .85 | 5 | 10001101 | .62 | 6 |
| $N = 6$ | $m = 3$ | | 1010011 | .41 | 5 | 10011011 | .85 | 5 | 01001110 | .62 | 6 |
| 000111 | 2.11 | 0 | 1001101 | .35 | 5 | | | | 01100101 | .57 | 7 |
| 001011 | 1.71 | 1 | 0101110 | .35 | 5 | 10100111 | .85 | 5 | | | |
| 010011 | 1.27 | 2 | | | | 01110011 | .80 | 6 | 01011001 | .57 | 7 |
| 001101 | 1.27 | 2 | | | | 01101101 | .72 | 6 | 10100011 | .38 | 7 |
| 010101 | .83 | 3 | 0111001 | .24 | 6 | 10101011 | .53 | 6 | 00111010 | .38 | 7 |
| | | | 1100011 | .00 | 6 | 01011110 | .47 | 6 | 10010101 | .32 | 7 |
| 100011 | .64 | 3 | 1010101 | .00 | 6 | | | | | | |

TABLE 1—*Continued*

| $R$ | $c_1$ | $U$ | $R$ | $c_1$ | $U$ | $R$ | $c_1$ | $U$ | $R$ | $c_1$ | $U$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 01010110 | .32 | 7 | 010000011 | 1.49 | 5 | 001011011 | 2.12 | 4 | 001111010 | .36 | 9 |
| | | | 001000101 | 1.49 | 5 | 000111101 | 2.06 | 4 | | | |
| 01101001 | .25 | 8 | 000101001 | 1.49 | 5 | | | | 100110101 | .30 | 9 |
| 10001110 | .05 | 7 | 000100110 | 1.23 | 5 | 010010111 | 2.06 | 4 | 010110101 | .30 | 9 |
| 11000011 | .00 | 8 | | | | 001101011 | 1.85 | 5 | 011100101 | .29 | 10 |
| 10100101 | .00 | 8 | 000110001 | 1.22 | 6 | 010100111 | 1.79 | 5 | 100011110 | .28 | 8 |
| 10011001 | .00 | 8 | 000011010 | 1.20 | 5 | 100001111 | 1.77 | 4 | 110001011 | .27 | 9 |
| | | | 001001001 | 1.19 | 6 | 001011101 | 1.76 | 5 | | | |
| 01100110 | .00 | 8 | 010000101 | 1.13 | 6 | | | | 101001101 | .27 | 9 |
| 01011010 | .00 | 8 | 100000011 | .93 | 6 | 010011011 | 1.76 | 5 | 011001110 | .27 | 9 |
| 00111100 | .00 | 8 | | | | 001110011 | 1.58 | 6 | 011011001 | .26 | 10 |
| $N = 9$ | $m = 2$ | | 001000110 | .93 | 6 | 000111110 | 1.50 | 5 | 101100011 | .09 | 10 |
| 001111111 | 2.42 | 0 | 000101010 | .93 | 6 | 100010111 | 1.50 | 5 | 100101110 | .01 | 9 |
| 010111111 | 2.06 | 1 | 001010001 | .92 | 7 | 011000111 | 1.49 | 6 | | | |
| 011011111 | 1.76 | 2 | 000011100 | .84 | 6 | | | | 110010011 | .00 | 10 |
| 100111111 | 1.50 | 2 | 010001001 | .83 | 7 | 001101101 | 1.49 | 6 | 001111100 | .00 | 10 |
| 011101111 | 1.49 | 3 | | | | 010101011 | 1.49 | 6 | 100111001 | .00 | 10 |
| | | | 000110010 | .66 | 7 | 010011101 | 1.40 | 6 | 101010101 | .00 | 10 |
| 101011111 | 1.20 | 3 | 001100001 | .65 | 8 | 100100111 | 1.23 | 6 | 0101110101 | .00 | 10 |
| 101101111 | .93 | 4 | 001001010 | .63 | 7 | 001110101 | 1.22 | 7 | | | |
| 011110111 | .92 | 4 | 100000101 | .57 | 7 | | | | 10011101 | .00 | 10 |
| 011111011 | .92 | 5 | 000101100 | .57 | 7 | 010110011 | 1.22 | 7 | $N = 10$ | $m = 2$ | |
| 110011111 | .84 | 4 | | | | 001011110 | 1.20 | 6 | 0011111111 | 2.54 | 0 |
| | | | 010000110 | .57 | 7 | 100011011 | 1.20 | 6 | 0101111111 | 2.20 | 1 |
| 101110111 | .66 | 5 | 010010001 | .56 | 8 | 011001011 | 1.19 | 7 | 0110111111 | 2.20 | 2 |
| 110101111 | .57 | 5 | 001010010 | .36 | 8 | 010101101 | 1.13 | 7 | 1001111111 | 1.66 | 2 |
| 011111101 | .56 | 6 | 000110100 | .30 | 8 | | | | 0111011111 | 1.66 | 3 |
| 101111011 | .36 | 6 | 010100001 | .29 | 9 | 101000111 | .93 | 7 | | | |
| 110110111 | .30 | 6 | | | | 001101110 | .93 | 7 | 0111101111 | 1.42 | 4 |
| | | | 100001001 | .27 | 8 | 100101011 | .93 | 7 | 1010111111 | 1.38 | 3 |
| 111001111 | .27 | 6 | 010001010 | .27 | 8 | 001111001 | .92 | 8 | 0111110111 | 1.16 | 5 |
| 111010111 | .00 | 7 | 001001100 | .27 | 8 | 011010011 | .92 | 8 | 1011011111 | 1.12 | 4 |
| 110111011 | .00 | 7 | 001100010 | .09 | 9 | | | | 1100111111 | 1.04 | 4 |
| 101111101 | .00 | 7 | 100000110 | .01 | 8 | 010110101 | .86 | 8 | | | |
| 011111110 | .00 | 7 | | | | 010011110 | .84 | 7 | 1011101111 | .88 | 5 |
| $N = 9$ | $m = 3$ | | 100010001 | .00 | 9 | 100011101 | .84 | 7 | 0111111011 | .88 | 6 |
| 000000111 | 2.99 | 0 | 010010010 | .00 | 9 | 011001101 | .83 | 8 | 1101011111 | .78 | 5 |
| 000001011 | 2.69 | 1 | 001010100 | .00 | 9 | 001110110 | .66 | 8 | 1011110111 | .62 | 6 |
| 000010011 | 2.42 | 2 | 000111000 | .00 | 9 | | | | 1101101111 | .54 | 6 |
| 000001101 | 2.33 | 2 | $N = 9$ | $m = 4$ | | 100110011 | .66 | 8 | | | |
| 000100011 | 2.15 | 3 | 000011111 | 3.26 | 0 | 011100011 | .65 | 9 | 0111111101 | .54 | 7 |
| | | | 000101111 | 2.99 | 1 | 101001011 | .63 | 8 | 1110011111 | .50 | 6 |
| 000010101 | 2.06 | 3 | 000110111 | 2.72 | 2 | 110000111 | .57 | 8 | 1011111011 | .34 | 7 |
| 001000011 | 1.85 | 4 | 001001111 | 2.69 | 2 | 100101101 | .57 | 8 | 1101110111 | .28 | 7 |
| 000100101 | 1.79 | 4 | 000111011 | 2.42 | 3 | | | | 1110101111 | .26 | 7 |
| 000001110 | 1.77 | 3 | | | | 010101110 | .57 | 8 | | | |
| 000011001 | 1.76 | 4 | 001010111 | 2.42 | 3 | 011010101 | .56 | 9 | 1111001111 | .00 | 8 |
| | | | 010001111 | 2.33 | 3 | 010111001 | .56 | 9 | 1110110111 | .00 | 8 |
| 000010110 | 1.50 | 4 | 001100111 | 2.15 | 4 | 101010011 | .36 | 9 | 1101111011 | .00 | 8 |
| | | | | | | | | | 1011111101 | .00 | 8 |

TABLE 1—*Continued*

| $R$ | $c_1$ | $U$ | $R$ | $c_1$ | $U$ | $R$ | $c_1$ | $U$ | $R$ | $c_1$ | $U$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0111111110 | .00 | 8 | 0101111110 | .66 | 8 | 0110001111 | 1.92 | 6 | 1010011011 | .84 | 9 |
| $N = 10$ | $m = 3$ | | 0111011101 | .66 | 9 | 0010111101 | 1.92 | 6 | 0101110101 | .82 | 10 |
| 0001111111 | 3.20 | 0 | 1011100111 | .50 | 9 | 0100111011 | 1.92 | 6 | 0110101101 | .80 | 10 |
| 0010111111 | 2.92 | 1 | 0111110011 | .50 | 10 | 0011101011 | 1.76 | 7 | | | |
| 0011011111 | 2.66 | 2 | 1011011011 | .46 | 9 | | | | 1100010111 | .78 | 9 |
| 0100111111 | 2.58 | 2 | * | | | 0101100111 | 1.70 | 7 | 1001011101 | .78 | 9 |
| 0011101111 | 2.42 | 3 | 0111101101 | .42 | 10 | 1001001111 | 1.66 | 6 | 0101011110 | .78 | 9 |
| | | | 1101010111 | .40 | 9 | 1000110111 | 1.66 | 6 | 1001011110 | .78 | 10 |
| 0101011111 | 2.32 | 3 | 1110001111 | .38 | 9 | 0001111110 | 1.66 | 6 | 1011000111 | .62 | 10 |
| 0011110111 | 2.16 | 4 | 1100111011 | .38 | 9 | 0110010111 | 1.66 | 7 | | | |
| 0101101111 | 2.08 | 4 | 1010111101 | .38 | 9 | | | | 1001110011 | .62 | 10 |
| 1000111111 | 2.04 | 3 | | | | 0101011011 | 1.66 | 7 | 0011110110 | .62 | 10 |
| 0110011111 | 2.04 | 4 | 0110111110 | .38 | 9 | 0011011101 | 1.66 | 7 | 0111010011 | .62 | 11 |
| | | | 1011101011 | .22 | 10 | 0100111101 | 1.58 | 7 | 1010101011 | .60 | 10 |
| 0011111011 | 1.88 | 5 | 1101100111 | .16 | 10 | 0011110011 | 1.50 | 8 | 1100100111 | .54 | 10 |
| 0101110111 | 1.82 | 5 | 0111110101 | .16 | 11 | 0110100111 | 1.42 | 8 | | | |
| 0110101111 | 1.80 | 5 | 1001111110 | .12 | 9 | | | | 1001101101 | .54 | 10 |
| 1001011111 | 1.78 | 4 | | | | 0101101011 | 1.42 | 8 | 0101101110 | .54 | 10 |
| 1001101111 | 1.54 | 5 | 1110010111 | .12 | 10 | 0011101101 | 1.42 | 8 | 0111001101 | .54 | 11 |
| | | | 1101011011 | .12 | 10 | 1001010111 | 1.40 | 7 | 0110110101 | .54 | 11 |
| 0111001111 | 1.54 | 6 | 1011011101 | .12 | 10 | 1010001111 | 1.38 | 7 | 0101111001 | .54 | 11 |
| 0110110111 | 1.54 | 6 | 0111011110 | .12 | 10 | 1000111011 | 1.38 | 7 | | | |
| 0101111011 | 1.54 | 6 | 1100111101 | .04 | 10 | | | | 1000111110 | .50 | 9 |
| 0011111101 | 1.54 | 6 | $N = 10$ | $m = 4$ | | 0010111110 | 1.38 | 7 | 1100011011 | .50 | 10 |
| 1010011111 | 1.50 | 5 | 0000111111 | 3.58 | 0 | 0110011011 | 1.38 | 8 | 1010011101 | .50 | 10 |
| | | | 0001011111 | 3.32 | 1 | 0101011101 | 1.32 | 8 | 0110011110 | .50 | 10 |
| 1001110111 | 1.28 | 6 | 0001101111 | 3.08 | 2 | 1001100111 | 1.16 | 8 | 0111100011 | .38 | 12 |
| 0111010111 | 1.28 | 7 | 0010011111 | 3.04 | 2 | 0111000111 | 1.16 | 9 | | | |
| 1010101111 | 1.26 | 6 | 0001110111 | 2.82 | 3 | | | | 1011001011 | .34 | 11 |
| 0110111011 | 1.26 | 7 | | | | 0101110011 | 1.16 | 9 | 1010110011 | .34 | 11 |
| 0101111101 | 1.20 | 7 | 0010101111 | 2.80 | 3 | 0011110101 | 1.16 | 9 | 0011111010 | .34 | 11 |
| | | | 0100011111 | 2.70 | 3 | 0110101011 | 1.14 | 9 | 1101000111 | .28 | 11 |
| 1100011111 | 1.16 | 6 | 0011001111 | 2.54 | 4 | 1010010111 | 1.12 | 8 | 1001110101 | .28 | 11 |
| 0111100111 | 1.04 | 8 | 0010110111 | 2.54 | 4 | 1001011011 | 1.12 | 8 | | | |
| 1011001111 | 1.00 | 7 | 0001111011 | 2.54 | 4 | | | | 0101110110 | .28 | 11 |
| 1010110111 | 1.00 | 7 | | | | 0011011110 | 1.12 | 8 | 0111010101 | .28 | 12 |
| 1001111011 | 1.00 | 7 | 0100101111 | 2.46 | 4 | 0101101101 | 1.08 | 9 | 1100101011 | .26 | 11 |
| | | | 0011010111 | 2.28 | 5 | 1100001111 | 1.04 | 8 | 1010101101 | .26 | 11 |
| 0011111110 | 1.00 | 7 | 0010111011 | 2.26 | 5 | 1000111101 | 1.04 | 8 | 0110101110 | .26 | 11 |
| 0111011011 | 1.00 | 8 | 0101001111 | 2.20 | 5 | 0100111110 | 1.04 | 8 | | | |
| 1100101111 | .92 | 7 | 0001111101 | 2.20 | 5 | | | | 0110111001 | .26 | 12 |
| 0110111101 | .92 | 8 | | | | 0110011101 | 1.04 | 9 | 1100011101 | .16 | 11 |
| 0111101011 | .76 | 9 | 0100110111 | 2.20 | 5 | 1010100111 | .88 | 9 | 1011010011 | .08 | 12 |
| | | | 1000011111 | 2.16 | 4 | 1001101011 | .88 | 9 | 0111100101 | .04 | 13 |
| 1011010111 | .74 | 8 | 0011100111 | 2.04 | 6 | 0011101110 | .88 | 9 | 1001101110 | .00 | 11 |
| 1010111011 | .72 | 8 | 0011011011 | 2.00 | 6 | 0111001011 | .88 | 10 | | | |
| 1101001111 | .66 | 8 | 0101010111 | 1.94 | 6 | | | | 1110000111 | .00 | 12 |
| 1100110111 | .66 | 8 | | | | 0110110011 | .88 | 10 | 1101001011 | .00 | 12 |
| 1001111101 | .66 | 8 | 1000101111 | 1.92 | 5 | 0011111001 | .88 | 10 | 1100110011 | .00 | 12 |

TABLE 1—*Continued*

| $R$ | $c_1$ | $U$ | $R$ | $c_1$ | $U$ | $R$ | $c_1$ | $U$ | $R$ | $c_1$ | $U$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1011001101 | .00 | 12 | 0010011101 | 2.04 | 6 | 0100011110 | 1.16 | 8 | 0011100110 | .50 | 11 |
| 1010110101 | .00 | 12 | 0011001011 | 1.88 | 7 | | | | 0011110001 | .50 | 12 |
| | | | 0010110011 | 1.88 | 7 | 0101100011 | 1.04 | 10 | 0111000011 | .50 | 12 |
| 1001111001 | .00 | 12 | 0101000111 | 1.82 | 7 | 0011100101 | 1.04 | 10 | | | |
| 0111001110 | .00 | 12 | 0001110101 | 1.82 | 7 | 1010000111 | 1.00 | 9 | 1010010011 | .46 | 11 |
| 0110110110 | .00 | 12 | | | | 1000110011 | 1.00 | 9 | 0011011010 | .46 | 11 |
| 0101111010 | .00 | 12 | 0100101011 | 1.80 | 7 | 1001001011 | 1.00 | 9 | 0110100101 | .42 | 12 |
| 0011111100 | .00 | 12 | 0010101101 | 1.80 | 7 | | | | 0101101001 | .42 | 12 |
| | | | 1000010111 | 1.78 | 6 | 0011001110 | 1.00 | 9 | 1001010101 | .40 | 11 |
| 0111011001 | .00 | 13 | 0001011110 | 1.78 | 6 | 0010110110 | 1.00 | 9 | | | |
| $N = 10$ | $m = 5$ | | 0100011101 | 1.70 | 7 | 0001111010 | 1.00 | 9 | 0101010110 | .40 | 11 |
| 0000011111 | 3.70 | 0 | | | | 0110010011 | 1.00 | 10 | 1000101110 | .38 | 10 |
| 0000101111 | 3.46 | 1 | 0011010011 | 1.62 | 8 | 0011011001 | 1.00 | 10 | 1100001011 | .38 | 11 |
| 0000110111 | 3.20 | 2 | 1000100111 | 1.54 | 7 | | | | 1010001101 | .38 | 11 |
| 0001001111 | 3.20 | 2 | 0001101110 | 1.54 | 7 | 0101010101 | .94 | 10 | 1000111001 | .38 | 11 |
| 0001010111 | 2.94 | 3 | 0110000111 | 1.54 | 8 | 0100101110 | .92 | 9 | | | |
| | | | 0101001011 | 1.54 | 8 | 1000101101 | .92 | 9 | 0110001110 | .38 | 11 |
| 0010001111 | 2.92 | 3 | | | | 0110001101 | .92 | 10 | 0100111010 | .38 | 11 |
| 0000111011 | 2.92 | 3 | 0100110011 | 1.54 | 8 | 0100111001 | .92 | 10 | 0010111100 | .38 | 11 |
| 0001100111 | 2.70 | 4 | 0011001101 | 1.54 | 8 | | | | 0110011001 | .38 | 12 |
| 0010010111 | 2.66 | 4 | 0010110101 | 1.54 | 8 | 0110100011 | .76 | 11 | 1010100011 | .22 | 12 |
| 0001011011 | 2.66 | 4 | 0001111001 | 1.54 | 8 | 0011101001 | .76 | 11 | | | |
| | | | 1000011011 | 1.50 | 7 | 1001010011 | .74 | 10 | 0011101010 | .22 | 12 |
| 0000111101 | 2.58 | 4 | | | | 0011010110 | .74 | 10 | 1001100101 | .16 | 12 |
| 0100001111 | 2.58 | 4 | 0010011110 | 1.50 | 7 | 1010001011 | .72 | 10 | 0101100110 | .16 | 12 |
| 0010100111 | 2.42 | 5 | 0100101101 | 1.46 | 8 | | | | 0111000101 | .16 | 13 |
| 0001101011 | 2.42 | 5 | 0011100011 | 1.38 | 9 | 0010111010 | .72 | 10 | 0101110001 | .16 | 13 |
| 0010011011 | 2.38 | 5 | 1001000111 | 1.28 | 8 | 0101100101 | .70 | 11 | | | |
| | | | 0001110110 | 1.28 | 8 | 1001001101 | .66 | 10 | 0110101001 | .14 | 13 |
| 0100010111 | 2.32 | 5 | | | | 1000110101 | .66 | 10 | 1001001110 | .12 | 11 |
| 0001011101 | 2.32 | 5 | 0101010011 | 1.28 | 9 | 1100000111 | .66 | 10 | 1000110110 | .12 | 11 |
| 0011000111 | 2.16 | 6 | 0011010101 | 1.28 | 9 | | | | 1100010011 | .12 | 12 |
| 0001110011 | 2.16 | 6 | 1000101011 | 1.26 | 8 | 0101001110 | .66 | 10 | 1010010101 | .12 | 12 |
| 0010101011 | 2.14 | 6 | 0010101110 | 1.26 | 8 | 0100110110 | .66 | 10 | | | |
| | | | 0010111001 | 1.26 | 9 | 0001111100 | .66 | 10 | 0110010110 | .12 | 12 |
| 0100100111 | 2.08 | 6 | | | | 0110010101 | .66 | 11 | 0101011010 | .12 | 12 |
| 0001101101 | 2.08 | 6 | 0110001011 | 1.26 | 9 | 0101011001 | .66 | 11 | 0011011100 | .12 | 12 |
| 0000111110 | 2.04 | 5 | 0101001101 | 1.20 | 9 | | | | 1001011001 | .12 | 12 |
| 1000001111 | 2.04 | 5 | 0100110101 | 1.20 | 9 | 1000011110 | .62 | 9 | 1100001101 | .04 | 12 |
| 0100011011 | 2.04 | 6 | 1000011101 | 1.16 | 8 | 1001100011 | .50 | 11 | 0100111100 | .04 | 12 |

**9. The $c_1(R)$ test compared with Mann and Whitney $U$ test.** The two best known rank order tests now in current use for testing whether two samples come from the same population are the Mann and Whitney $U$ test [6] and the Wilcoxon $T$ test [16]. These two tests have been developed to test the null hypothesis that the sample $(x_1, \cdots, x_m)$ comes from the same population as the sample $(y_1, \cdots, y_n)$ with cdf $F(\xi)$ against the alternative that the cdf of $(y_1, \cdots, y_n)$,

TABLE 2

*Critical values of $c_1(R)$*

Listed below are the critical values of $c_1(R)$ tabulated as $c_1$ for the two-sample case for values of $N$, $m$ such that $6 \leqq N \leqq 10$, $1 < m \leqq [N/2]$, together with the probability ($p$) of exceeding this value on the null hypothesis, where $p \leqq .10$.

| $N$ | $m$ | $c_1$ | $p$ | $N$ | $m$ | $c_1$ | $p$ | $N$ | $m$ | $c_1$ | $p$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 6 | 3 | 2.11 | .050 | 9 | 2 | 2.42 | .028 | 10 | 3 | 3.20 | .008 |
| | | 1.71 | .100 | | | 2.06 | .056 | | | 2.92 | .016 |
| | | | | | | 1.76 | .083 | | | 2.66 | .025 |
| | | | | | | | | | | 2.58 | .033 |
| 7 | 2 | 2.11 | .048 | 9 | 3 | 2.99 | .012 | | | 2.42 | .041 |
| | | 1.70 | .095 | | | 2.69 | .023 | | | 2.32 | .050 |
| | | | | | | 2.42 | .035 | | | 2.16 | .058 |
| 7 | 3 | 2.46 | .029 | | | 2.33 | .048 | | | 2.08 | .067 |
| | | 2.11 | .057 | | | 2.06 | .071 | | | 2.04 | .083 |
| | | 1.76 | .086 | | | 1.85 | .083 | | | 1.88 | .091 |
| | | | | | | 1.79 | .095 | | | 1.82 | .100 |
| 8 | 2 | 2.27 | .036 | 9 | 4 | 3.26 | .008 | 10 | 4 | 3.32 | .010 |
| | | 1.89 | .071 | | | 2.99 | .016 | | | 3.04 | .019 |
| | | | | | | 2.69 | .032 | | | 2.80 | .029 |
| 8 | 3 | 2.75 | .018 | | | 2.42 | .048 | | | 2.70 | .033 |
| | | 2.42 | .036 | | | 2.33 | .056 | | | 2.54 | .048 |
| | | 2.12 | .054 | | | 2.15 | .063 | | | 2.28 | .057 |
| | | 2.04 | .071 | | | 2.12 | .071 | | | 2.20 | .076 |
| | | 1.80 | .089 | | | 2.06 | .087 | | | 2.04 | .086 |
| | | | | | | 1.85 | .095 | | | 1.94 | .095 |
| 8 | 4 | 2.89 | .014 | | | | | | | | |
| | | 2.59 | .029 | 10 | 2 | 2.54 | .022 | 10 | 5 | 3.46 | .008 |
| | | 2.27 | .057 | | | 2.20 | .067 | | | 2.92 | .028 |
| | | 1.95 | .071 | | | | | | | 2.66 | .040 |
| | | 1.89 | .100 | | | | | | | 2.58 | .048 |
| | | | | | | | | | | 2.32 | .067 |
| | | | | | | | | | | 2.14 | .079 |
| | | | | | | | | | | 2.08 | .087 |

$G(\xi)$ have the relation $F(\xi) > G(\xi)$ for every $\xi$. The $U$ test has been shown to be equivalent to the Wilcoxon $T$ test [6]. In fact Mann and Whitney show the following linear relation between the $U$ statistic and the $T$ statistic:

$$U = mn + \tfrac{1}{2}(n)(n+1) - T. \tag{9.1}$$

The $U$ test has been defined by Mann and Whitney as follows. In an ordered sample of $m$ $x$'s and $n$ $y$'s, let $U$ count the number of times a $y$ precedes an $x$. If $P(U < \bar{U})$ equals $\alpha$ under the null hypothesis, the test will be considered significant at the significance level $\alpha$ if $U < \bar{U}$ and the hypothesis of identical distribution of $x$ and $y$ will be rejected. Wilcoxon's $T$ is the sum of the ranks in the sample of $n$ observations.

The $c_1(R)$ test is generally more sensitive than the $U$ test or the equivalent $T$ test, for in many permutations where $U(R_i) = U(R_j)$, $c_1(R_i) \neq c_1(R_j)$, $i \neq j$, with the result that although

$$P\{U \leqq U(R_i)\} = P\{U \leqq U(R_j)\},$$
$$P\{c_1(R) \geqq c_1(R_i)\} \neq P\{c_1(R) \geqq c_1(R_j)\}, \qquad i \neq j.$$

The most powerful test is not difficult to compute, as it involves at most $n$ additions or subtractions of numbers readily accessible in the Fisher and Yates tables [3]. Table XX tabulates for $N \leqq 50$ the positive values of $EZ_{Ni}$.

There is a very close dependence between these two statistics, as a scrutiny of Table 1 will show. For all $N \leqq 8$, $m \geqq 1$, $U$ is a nonincreasing function of $c_1$, and for all $N$, $m = 1$, $U$ is a strictly decreasing function of $c_1$. In general, however, there is no such functional relationship. We shall now derive the correlation under $H_0$ between the $U$ and the $c_1$ statistic.

By definition [3], the correlation $\rho[c_1(R), U]$ can be written

$$\rho[c_1(R), U] = \frac{E_0[c_1(R)U]}{\sqrt{V_0(c_1(R)V(U))}}. \tag{9.2}$$

Since there is a linear relation [6] between the Mann and Whitney $U$ and Wilcoxon's $T$, to wit: $U = mn + \frac{1}{2}m(n+1) - T$, it follows that

$$V_0(U) = V_0(T), \qquad E_0(c_1(R)U) = -E_0(c_1(R)T), \tag{9.3}$$

where $T$ is the sum of the ranks of the sample of $n$ observations. We shall use Wilcoxon's statistic to evaluate $E_0(c_1(R)U)$. Now

$$E_0(c_1(R)T) = \frac{1}{{}_NC_n} \sum \left( \sum_{i=m+1}^{m+n} EZ_{Nr_i} \right) \left( \sum_{j=m+1}^{m+n} r_j \right), \tag{9.4}$$

where $\sum$ denotes the summation over the ${}_NC_n$ distinct combinations. Consider an $r_j = k$ fixed. Then the coefficients, $EZ_{Nr_i}$, of $k$ in the summation occur ${}_{N-1}C_{n-1}$ times if $r_i = k$, and ${}_{N-2}C_{n-2}$ times if $r_i \neq k$. Since $\sum_{i=1}^{N} kEZ_{Nr_i} = 0$, the coefficient of $k$ reduces to $({}_{N-1}C_{n-1} - {}_{N-2}C_{n-2})EZ_{Nk}$. Hence it follows that

$$\begin{aligned} E(c_1(R)T) &= \frac{1}{{}_NC_n} ({}_{N-1}C_{n-1} - {}_{N-2}C_{n-2}) \sum_{k=1}^{N} kE(Z_{Nk}) \\ &= \frac{mn}{N(N-1)} \sum_{k=1}^{N} kEZ_{Nk}. \end{aligned} \tag{9.5}$$

Now

$$\sum_{i=1}^{N} iEZ_{Ni} = \sum_{i=1}^{N} iN\binom{N-1}{i-1}\int_{-\infty}^{\infty} xf(x)[F(x)]^{i-1}[1-F(x)]^{N-1}\,dx$$

$$= N(N-1)\int_{-\infty}^{\infty} xf(x)F(x)\,dx, \tag{9.6}$$

after interchanging the order of summation and integration and using the well known properties of a binomial. Integrating the right-hand side by parts we have

$$\sum_{i=1}^{N} iEZ_{Ni} = N(N-1)\int_{-\infty}^{\infty}[f(x)]^2\,dx = \frac{N(N-1)}{2\sqrt{\pi}}. \tag{9.7}$$

Now

$$V_0(U) = V_0(T) = \frac{mn(N+1)}{12}, \tag{9.8}$$

[6] and

$$V_0(c_1(R)) = \frac{mn\mu_2}{N-1}, \tag{9.9}$$

where $\mu_2 = \sum(EZ_{Ni})^2/N$. Since $E_0c_1(R) = 0$, the covariance of $c_1(R)$ and $T$ is equal to $E_0(c_1(R)T)$. Hence

$$\rho = -\frac{\frac{mn}{2\sqrt{\pi}}}{mn\sqrt{\frac{N+1}{12(N-1)}\mu_2}} = -\sqrt{\frac{3(N-1)}{\pi(N+1)\mu_2}}. \tag{9.10}$$

It has been shown in an unpublished paper of W. Hoeffding that

$$\lim_{N\to\infty}\sum_{i=1}^{N}(EZ_{Ni})^2/N = 1, \text{ and hence } \lim_{N\to\infty}\rho = -\sqrt{3/\pi} \cong -0.9772050.$$

Thus there is not a linear functional relationship between these two statistics. We do, however, note that, for any given $N$, $\rho$ is independent of the subdivision of $N$ into groups of $m$ and $n$, $m + n = N$.

Numerical values of $\rho$ can then easily be computed using Tables XX and XXI of Fisher and Yates [3]. From numerical calculation, $|\rho| > .99$, $N \leqq 15$.

**10. The power of the $c_1(R)$ test.** Since no method has yet been found for evaluating analytically the power function of the $c_1(R)$ test, the power of the $c_1(R)$ test has been investigated by drawing two random groups of samples of size $N = 8$, $m = 4$, $n = 4$ for $\delta = 0, .1, .2, .3, .4, .5, .6, .7, .8, .9, 1.0, 1.2, 1.5, 2, 2.5, 3, 4$. The samples were drawn by selecting 100 groups of eight pairs of digits except the pair 00 from the Fisher and Yates tables of random numbers [3]. These pairs were uniformly distributed over the range 1–99. These 800 pairs of two-digit numbers were then transformed into standard normal deviates $x$

using the relationship $F(x) = \int_{-\infty}^{x} f(x)\,dx$, where $F(x) = (.01)\ (N_i)$, $N_i = 1, \cdots, 99$.

The first four pairs in each group became the $y$ sample and the last four pairs became the $x$ sample. Then a constant, $\delta$, was added to each value of $y$ and the $c_1(R)$ test carried out. The results were plotted as number of times the $H_a$ (Section 5) was accepted per 100 against $\delta$, and a smooth curve fitted to the pairs of points.

Neyman and Tokarska [8] give tables of the power function of the one-tailed Student's $t$ test for levels of significance $\alpha = .05$ and .01. Although the sampling procedure yields the approximate power function for $\alpha = .057$, and the Neyman and Tokarska power function is given for $\alpha = .05$, general comparisons can be made on the evidence of shape. No attempt is made to compare actual values.

We can place confidence limits on $\beta(\delta)$, the actual value of the power, and evaluate the reliability of the sampling estimates of $\beta(\delta)$. Let $M$ be the total number of observations and $s$ the number of observations that lead to rejection. Then $s$ has a binomial distribution with parameters $M$ and $\beta(\delta)$, and $E(s) = M\beta((\delta)$; $V(S) = M\beta(\delta)(1 - \beta(\delta))$.

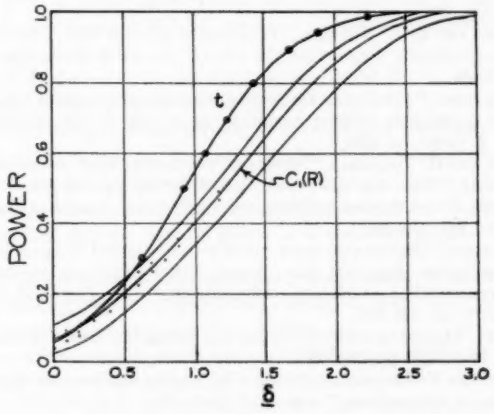



FIG. 1

For large $M$, $(s - M\beta(\delta))/\sqrt{\beta(\delta)(1 - \beta(\delta))M} \sim \dot{M}(0, 1)$ and we have as confidence limits on $\beta(\delta)$ at the level $\alpha$ [1]

$$\frac{s + \frac{1}{2}t_a^2 - t_a\sqrt{s\left(1 - \frac{s}{M}\right) + \frac{1}{4}t_a^2}}{M + t_a^2} \leqq \beta(\delta) \leqq \frac{s + \frac{1}{2}t_a^2 + t_a\sqrt{s\left(1 - \frac{s}{M}\right) + \frac{1}{4}t_a^2}}{M + t_a^2}. \tag{10.1}$$

Figure 1 shows the power curve of the one-tailed $t$-test and that of the $c_1$ test; 95% confidence bands for the true power of the $c_1$ test are shown.

Curve $t$ is the power curve of the one-tailed $t$-test with 6 d.f. such that $\beta(0) = .05$. Curve $c_1(R)$ is the power curve of the $c_1(R)$ test with $N = 8$, $m = 4$, $n = 4$ at the .057 level, together with the confidence band on the true power of the $c_1(R)$.

**Acknowledgement.** The author wishes to acknowledge the valuable help of Professor W. Hoeffding who suggested the topic and under whose direction the work was completed.

## REFERENCES

[1] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, 1946.

[2] W. J. Dixon, "A criterion for testing the hypothesis that two samples are from the same population," *Annals of Math. Stat.*, Vol. 11 (1940), pp. 199–204.

[3] R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research*, 3rd ed., Hafner, 1949.

[4] W. Hoeffding, " 'Optimum' nonparametric tests," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1951.

[5] M. G. Kendall, *Advanced Theory of Statistics*, Vol. I and II, Charles Griffin and Co., London, 1946.

[6] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *Annals of Math. Stat.*, Vol. 18 (1947), pp. 50–60.

[7] K. K. Matthen, "A criterion for testing whether two samples have come from the same population without assuming the nature of the population," *Sankhyā*, Vol. 7 (1946), p. 329.

[8] J. Neyman and B. Tokarska, "Errors of the second kind in testing 'Student's' hypothesis," *Jour. Am. Stat. Assn.*, Vol. 31 (1936), pp. 318–326.

[9] G. Noether, "On a theorem of Wald and Wolfowitz," *Annals of Math. Stat.*, Vol. 20 (1949), pp. 455–458.

[10] E. J. G. Pitman, "Significance tests which may be applied to samples from any population," *Jour. Roy. Stat. Soc., Suppl.*, Vol. 4 (1937), pp. 119–130.

[11] E. J. G. Pitman, "The correlation coefficient test," *Jour. Roy. Stat. Soc., Suppl.*, Vol. 4 (1937), pp. 225–232.

[12] H. Scheffé, "Statistical inference in the non-parametric case," *Annals of Math. Stat.*, Vol. 14 (1943), pp. 305–330.

[13] F. S. Swed and C. Eisenhart, "Tables for testing randomness of grouping in a sequence of alternatives," *Annals of Math. Stat.*, Vol. 14 (1943), pp. 66–87.

[14] A. Wald and J. Wolfowitz, "On a test whether two samples are from the same population," *Annals of Math. Stat.*, Vol. 11 (1940), pp. 147–162.

[15] A. Wald and J. Wolfowitz, "Statistical tests based on permutations of the observations," *Annals of Math. Stat.*, Vol. 15 (1944), pp. 358–372.

[16] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bull.*, Vol. 1 (1945), pp. 80–83.

[17] J. Wolfowitz, "Additive partition functions and a class of statistical hypotheses," *Annals of Math. Stat.*, Vol. 13 (1942), pp. 247–279.

# COMBINATORIAL PROPERTIES OF GROUP DIVISIBLE INCOMPLETE BLOCK DESIGNS

By R. C. Bose and W. S. Connor[1]

*University of North Carolina*

**1. Summary and introduction.** Group divisible incomplete block designs are an important subclass of partially balanced designs [1], [2] with two associate classes, and they may also be regarded as a special case of intra- and inter-group balanced incomplete block designs [3], [4]. They may be defined as follows.

An incomplete block design with $v$ treatments each replicated $r$ times in $b$ blocks of size $k$ is said to be group divisible (GD) if the treatments can be divided into $m$ groups, each with $n$ treatments, so that the treatments belonging to the same group occur together in $\lambda_1$ blocks and treatments belonging to different groups occur together in $\lambda_2$ blocks. If $\lambda_1 = \lambda_2 = \lambda$ (say), then every pair of treatments occurs in $\lambda$ blocks, and the design becomes a balanced incomplete block design, which has been extensively studied [5], [6], [7], [8]. We shall therefore confine ourselves to the case $\lambda_1 \neq \lambda_2$.

The object of this paper is to study the combinatorial properties of these designs. It is shown that the GD designs can be divided into three exhaustive and mutually exclusive classes:

(a) Singular GD designs characterized by $r - \lambda_1 = 0$;

(b) Semi-regular GD designs characterized by $r - \lambda_1 > 0$, $rk - v\lambda_2 = 0$;

(c) Regular GD designs characterized by $r - \lambda_1 > 0$, $rk - v\lambda_2 > 0$.

Certain inequality relations between the parameters necessary for the existence of the design have been derived in each case. Some other interesting theorems about the structure of these designs have also been obtained. Methods of constructing GD designs will be given in a separate paper.

**2. Group divisible designs regarded as special cases of partially balanced and intra- and inter-group incomplete block designs.** A partially balanced design with two associate classes is defined as follows.

(i) There are $v$ treatments, each replicated $r$ times in $b$ blocks of size $k$.

(ii) There can be established a relation of association between any two treatments satisfying the following conditions. (a) Two treatments are either first associates or second associates. (b) Each treatment has $n_1$ first associates and $n_2$ second associates. (c) Given any two treatments which are $i$th associates, the number of treatments which is common to the $j$th associates of the first and $k$th associates of the second is $p_{jk}^{i}$ and is independent of the pair of treatments with which we start ($i, j, k = 1, 2$). Also $p_{jk}^{i} = p_{ik}^{j}$.

(iii) Two treatments which are $i$th associates occur together in exactly $\lambda_i$ blocks ($i = 1, 2$).

[1] Now with the National Bureau of Standards, Washington, D. C.

It is known [1] that the following relations hold among the parameters:

$$bk = vr,\ v = n_1 + n_2 + 1,\ n_1\lambda_1 + n_2\lambda_2 = r(k-1), \tag{2.0}$$

$$1 + p_{11}^1 + p_{12}^1 = p_{11}^2 + p_{12}^2 = n_1, \tag{2.1}$$

$$p_{21}^1 + p_{22}^1 = 1 + p_{21}^2 + p_{22}^2 = n_2, \tag{2.2}$$

$$n_1 p_{12}^1 = n_2 p_{11}^2,\ n_1 p_{22}^1 = n_2 p_{12}^2. \tag{2.3}$$

If in a GD design with parameters $v, b, r, k, \lambda_1, \lambda_2, m, n$, treatments belonging to the same group are considered as first associates, and treatments belonging to different groups are considered as second associates, then it is easy to see that it is a partially balanced design with two associate classes for which

$$n_1 = n - 1, \qquad n_2 = n(m-1), \tag{2.4}$$

$$p_{11}^1 = n - 2, \qquad p_{12}^1 = 0, \qquad p_{22}^1 = n(m-1), \tag{2.5}$$

$$p_{11}^2 = 0, \qquad p_{12}^2 = n - 1, \qquad p_{22}^2 = n(m-2). \tag{2.6}$$

Conversely suppose for a partially balanced design $p_{12}^1 = 0$. Then from (2.1), $p_{11}^1 = n_1 - 1$. The relation of first association between two treatments is by definition commutative. We shall show that in the present case it is also transitive. Let the treatments $\theta_0$ and $\theta_1$ be first associates. Let the other first associates of $\theta_0$ be $\theta_2, \theta_3, \cdots, \theta_{n_1}$. Now since $\theta_0$ and $\theta_1$ have $n_1 - 1$ common first associates, they can be no other than $\theta_2, \theta_2, \cdots, \theta_{n_1}$. Also since $\theta_1$ has exactly $n_1$ first associates, so all its first associates are $\theta_0, \theta_2, \theta_3, \cdots, \theta_{n_1}$. This shows that any first associate of $\theta_0$ (other than $\theta_1$) is also a first associate of $\theta_1$. These conditions are sufficient to insure that the $v$ treatments can be divided into groups of $n_1 + 1$ such that two treatments in the same group are first associates, and two treatments in different groups are second associates. Hence the design is a GD design where the treatments of the same group are first associates. Similarly if $p^2{}_{12} = 0$, we can show that the partially balanced design is a GD design, the treatments of the same group being second associates. We can therefore state

THEOREM 1. *The necessary and sufficient condition for a partially balanced design to be group divisible is the vanishing of $p_{12}^1$ or $p_{12}^2$. If $p_{12}^i = 0$ then the treatments in the same group are $i$th associates* $(i = 1, 2)$.

A GD design is also a special case of intra- and inter-group balanced incomplete block designs which have been defined by Nair and Rao [3] in the following manner.

(i) there are $v$ treatments arranged in $b$ blocks of size $k$.

(ii) the $v$ treatments fall into $m$ groups consisting of $v_1, v_2, \cdots, v_m$ treatments, treatments of the $i$th group being replicated $r_i$ times $(i = 1, 2, \cdots, m)$.

(iii) every pair of treatments of the $i$th group occurs in $\lambda_{ii}$ blocks $(i = 1, 2, \cdots, m)$, and every pair of treatments belonging to the $i$th and $j$th groups, $i \neq j$, occurs in $\lambda_{ij}$ blocks $(i, j = 1, 2, \cdots, m)$.

If in particular

$$r_i = r, \qquad \lambda_{ii} = \lambda_1 \qquad (i = 1, 2, \cdots, m),$$
$$\lambda_{ij} = \lambda_2 \qquad (i \neq j;\ i, j = 1, 2, \cdots, m),$$

then the design reduces to a GD design.

GD designs may therefore be characterized as the class of designs which are partially balanced and intra- and inter-group balanced at the same time.

**3. Relations between the parameters, and the classification of GD designs.** Clearly

$$v = mn, \qquad bk = vr. \tag{3.0}$$

Any given treatment $\theta$ occurs in $r$ blocks. Since each of these blocks contains $k - 1$ other treatments, there are $r(k - 1)$ pairs of which one member is $\theta$. But $\theta$ must form $\lambda_1$ pairs with each of the $n - 1$ treatments belonging to the same group as $\theta$, and $\lambda_2$ pairs with each of the $n(m - 1)$ treatments not in the same group as $\theta$. Hence

$$(n - 1)\lambda_1 + n(m - 1)\lambda_2 = r(k - 1). \tag{3.1}$$

Also

$$r \geqq \lambda_1, \qquad r \geqq \lambda_2. \tag{3.2}$$

The eight parameters $v$, $b$, $r$, $k$, $\lambda_1$, $\lambda_2$, $m$, $n$ are therefore connected by the three relations (3.0) and (3.1), so that only five parameters are free.

Let $n_{ij} = 1$ or 0 according as the $i$th treatment does or does not occur in the $j$th block. Then the matrix

$$N = (n_{ij}) \tag{3.3}$$

is defined to be the incidence matrix of the design. From the conditions satisfied by the design, it is easy to see that

$$\sum_{j=1}^{v} n_{ij}^2 = r, \tag{3.4a}$$

$$\sum_{j=1}^{v} n_{ij} n_{uj} = \lambda_1 \text{ or } \lambda_2, \tag{3.4b}$$

according as the $i$th and $u$th treatments ($i \neq u$) do belong or do not belong to the same group.

If in numbering the treatments we follow the convention that the $l$th group consists of the treatments number $n(l - 1) + 1, n(l - 1) + 2, \cdots, nl$, then from (3.4) and (3.5) we can write

$$NN' = \begin{bmatrix} A & B & \cdots & B \\ B & A & \cdots & B \\ \vdots & \vdots & & \vdots \\ B & B & \cdots & A \end{bmatrix} \tag{3.5}$$

where $N'$ is the transpose of the matrix $N$, and $A$ and $B$ are $n \times n$ matrices defined by

$$A = \begin{bmatrix} r & \lambda_1 & \cdots & \lambda_1 \\ \lambda_1 & r & \cdots & \lambda_1 \\ \vdots & \vdots & & \vdots \\ \lambda_1 & \lambda_1 & \cdots & r \end{bmatrix}, \qquad B = \begin{bmatrix} \lambda_2 & \lambda_2 & \cdots & \lambda_2 \\ \lambda_2 & \lambda_2 & \cdots & \lambda_2 \\ \vdots & \vdots & & \vdots \\ \lambda_2 & \lambda_2 & \cdots & \lambda_2 \end{bmatrix}. \tag{3.6}$$

Each row or column in the matrix on the right-hand side in (3.5) contains $A$ in the diagonal position, and contains $B$ in the other $m - 1$ positions.

To evaluate $|NN'|$ proceed as follows.

(i) Add the 2nd, $\cdots$, $mn$th rows in $|NN'|$ to the first row, every element of which now becomes $r + (n-1)\lambda_1 + n(m-1)\lambda_2 = rk$ (cf. (3.1)).

(ii) Take $rk$ outside the determinant, and subtract the first row multiplied by $\lambda_2$ from all other rows. Then

$$|NN'| = rk \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 - \lambda_2 & r - \lambda_2 & \cdots & \lambda_1 - \lambda_2 \\ \vdots & \vdots & & \vdots \\ \lambda_1 - \lambda_2 & \lambda_1 - \lambda_2 & \cdots & r - \lambda_2 \end{vmatrix} \cdot \begin{vmatrix} r - \lambda_2 & \lambda_1 - \lambda_2 & \cdots & \lambda_1 - \lambda_2 \\ \lambda_1 - \lambda_2 & r - \lambda_2 & \cdots & \lambda_1 - \lambda_2 \\ \vdots & \vdots & & \vdots \\ \lambda_1 - \lambda_2 & \lambda_1 - \lambda_2 & \cdots & r - \lambda_2 \end{vmatrix}^{m-1}$$

$$= rk\,\{(r - \lambda_2) + (n-1)(\lambda_1 - \lambda_2)\}^{m-1} \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 - \lambda_2 & r - \lambda_2 & \cdots & \lambda_1 - \lambda_2 \\ \vdots & \vdots & & \vdots \\ \lambda_1 - \lambda_2 & \lambda_1 - \lambda_2 & \cdots & r_1 - \lambda_2 \end{vmatrix}^{m}.$$

Using (3.0) and (3.1) we can finally write

$$|NN'| = rk(rk - v\lambda_2)^{m-1}(r - \lambda_1)^{m(n-1)}. \tag{3.7}$$

The quantity $rk - v\lambda_2$ occurring above is nonnegative. Here we shall prove this statement for the case $r - \lambda_1 > 0$, and complete the proof later in Section 4.

Let $N_1$ be the submatrix formed from the matrix $N$ given by (3.3), by taking the first $2n$ rows (which correspond to the treatments of the first two groups).

Then

$$N_1 N_1' = \begin{bmatrix} A & B \\ B & A \end{bmatrix}. \tag{3.8}$$

$$| N_1 N_1' | = \{ r + (n-1)\lambda_1 + n\lambda_2 \} (r - \lambda_1)^{2(n-1)} (rk - v\lambda_2). \tag{3.9}$$

Since $| N_1 N_1' | \geqq 0$, it follows that $rk - v\lambda_2 \geqq 0$ if $r - \lambda_1 > 0$.

Hence we can divide all GD designs into three exhaustive and mutually exclusive classes:

(a) Singular GD designs characterized by $r = \lambda_1$;

(b) Semi-regular GD designs characterized by $r > \lambda_1$, $rk - v\lambda_2 = 0$;

(c) Regular GD designs characterized by $r > \lambda_1$, $rk - v\lambda_2 > 0$.

The main combinatorial properties of each of these three classes will be given in the succeeding sections.

**4. Singular GD designs.** Consider a balanced incomplete block design with $v^*$ treatments, each replicated $r^*$ times in $b^*$ blocks of size $k^*$, such that any two treatments occur together in $\lambda^*$ blocks. Replace each treatment with a group of $n$ treatments. Now there are $v = nv^*$ treatments divided into $v^*$ groups (each group corresponding to one of the original treatments). Two treatments belonging to the same group now occur together $r^*$ times and two treatments belonging to different groups occur together $\lambda^*$ times. We thus get a GD design with parameters

$$\begin{aligned} &v = nv^*, \quad b = b^*, \quad r = r^*, \quad k = nk^*, \\ &\lambda_1 = r^*, \quad \lambda_2 = \lambda^*, \quad m = v^*, \quad n = n, \end{aligned} \tag{4.0}$$

which belongs to the singular class since $r - \lambda_1 = 0$.

Conversely consider a singular GD design with parameters $v$, $b$, $r$, $k$, $\lambda_1$, $\lambda_2$, $m$, $n$, where $r = \lambda_1$. Let $\theta$ and $\phi$ be any two treatments belonging to the same group. $\theta$ occurs in $r$ blocks, and since $r = \lambda_1$, $\phi$ must occur in each of these $r$ blocks and nowhere else. Hence if a treatment occurs in a certain block, every treatment belonging to the group occurs in that block. Let each group of treatments be replaced by a single treatment in the design, then there are $v^* = m$ treatments in the new design, and because any two treatments belonging to different groups occur together $\lambda_2$ times in the original design, the new design is a balanced incomplete block design with parameters

$$v^* = m, \quad b^* = b, \quad r^* = r, \quad k^* = k/n, \quad \lambda^* = \lambda_2. \tag{4.1}$$

We may therefore state

THEOREM 2. *If in a balanced incomplete block design with parameters $v^*$, $b^*$, $r^*$, $k^*$, $\lambda^*$ each treatment is replaced by a group of $n$ treatments, we get a singular GD design with parameters given by* (4.0). *Conversely, every singular GD design is obtainable in this way from a corresponding balanced incomplete block design.*

COROLLARY. *For a singular GD design* $b \geqq m$.

For example, consider the balanced incomplete block design with parameters $v^* = b^* = 7, r^* = k^* = 3, \lambda^* = 1$. The plan for this is given below, the columns representing the blocks.

| | | | | | | |
|---|---|---|---|---|---|---|
| $a$ | $b$ | $c$ | $d$ | $e$ | $f$ | $g$ |
| $b$ | $c$ | $d$ | $e$ | $f$ | $g$ | $a$ |
| $d$ | $e$ | $f$ | $g$ | $a$ | $b$ | $c$ |

If $n = 2$, then we may replace the treatment $a$ by $a_1$, $a_2$, and do the same for the other treatments. We then get the singular GD design with the parameters

$$v = 14, \quad b = 7, \quad r = 3, \quad k = 6,$$

$$\lambda_1 = 3, \quad \lambda_2 = 1, \quad m = 7, \quad n = 2,$$

the plan for which is shown below.

| | | | | | | |
|---|---|---|---|---|---|---|
| $a_1$ | $b_1$ | $c_1$ | $d_1$ | $e_1$ | $f_1$ | $g_1$ |
| $a_2$ | $b_2$ | $c_2$ | $d_2$ | $e_2$ | $f_2$ | $g_2$ |
| $b_1$ | $c_1$ | $d_1$ | $e_1$ | $f_1$ | $g_1$ | $a_1$ |
| $b_2$ | $c_2$ | $d_2$ | $e_2$ | $f_2$ | $g_2$ | $a_2$ |
| $d_1$ | $e_1$ | $f_1$ | $g_1$ | $a_1$ | $b_1$ | $c_1$ |
| $d_2$ | $e_2$ | $f_2$ | $g_2$ | $a_2$ | $b_2$ | $c_2$ |

As before the columns represent the blocks.

The relation $rk - v\lambda_2 \geqq 0$ is true by definition for semi-regular and regular GD designs. We shall show that it holds for singular GD designs also.

Let the parameters of a singular GD design be given by (4.0). Then remembering the relation $\lambda^*(v^* - 1) = r^*(k^* - 1)$ which holds for a balanced incomplete block design

$$\begin{aligned} rk - v\lambda_2 &= n(r^*k^* - v^*\lambda^*) \\ &= n(r^* - \lambda^*) \\ &\geqq 0. \end{aligned}$$

Hence we may state

THEOREM 3. *For any GD design* $rk - v\lambda_2 \geqq 0$.

**5. Semi-regular GD designs.** For a semi-regular GD design we have by definition

$$r - \lambda_1 > 0, \qquad rk - v\lambda_2 = 0. \tag{5.0}$$

Hence

$$r + (n - 1)\lambda_1 = n\lambda_2. \tag{5.1}$$

We shall now prove the following

THEOREM 4. *For a semi-regular GD design $k$ is divisible by $m$. If $k = cm$ then every block must contain $c$ treatments from every group.*

Let $e_j$ treatments from the first group occur in the $j$th block ($j = 1, 2, \cdots, b$). Then

$$\sum_{j=1}^{b} e_j = nr, \tag{5.2}$$

$$\sum_{j=1}^{b} e_j(e_j - 1) = n(n-1)\lambda_1, \tag{5.3}$$

since each treatment from the first group occurs in $r$ blocks, and every pair of treatments from the first group occurs in $\lambda_1$ blocks. Using (5.1), (5.2) and (5.3),

$$\sum_{j=1}^{b} e_j^2 = n^2\lambda_2.$$

Let

$$\bar{e} = \frac{1}{b}\sum_{j=1}^{b} e_j = \frac{nr}{b} = \frac{k}{m}$$

from (3.0). Hence

$$\sum_{j=1}^{b} (e_j - \bar{e})^2 = n^2\lambda_2 - \frac{bk^2}{m^2} = 0 \tag{5.4}$$

from (3.0) and (5.0). Therefore

$$e_1 = e_2 = \cdots = e_b = \bar{e} = \frac{k}{m}. \tag{5.5}$$

Since $e_j$ must be integral, $k$ must be divisible by $m$. If $k = cm$ then $e_j = c$ ($j = 1, 2, \cdots, b$). The same argument applies to treatments of any other group. This proves our theorem.

The relation (3.7) shows that the $v \times v$ matrix $NN'$ given by (3.5) is singular, for the case of semi-regular GD designs. We shall now show that its rank is not less than $v - m + 1$. Without changing the rank of $NN'$ we can transform it into $M_1$ and then into $M_2$, where

$$M_1 = \begin{bmatrix} D & D & \cdots & D \\ B & A & \cdots & B \\ \vdots & \vdots & & \vdots \\ B & B & \cdots & A \end{bmatrix},$$

where $D = A + (m-1)B$, and

$$M_2 = \begin{bmatrix} D & 0 & \cdots & 0 \\ B & A-B & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ B & 0 & \cdots & A-B \end{bmatrix},$$

where 0 is the null $n \times n$ matrix.

If we strike out the last row and column from $A - B$ we get the $(n-1) \times (n-1)$ matrix $C$ given by

$$C = \begin{bmatrix} r-\lambda_2 & \lambda_1-\lambda_2 & \cdots & \lambda_1-\lambda_2 \\ \lambda_1-\lambda_2 & r-\lambda_2 & \cdots & \lambda_1-\lambda_2 \\ \vdots & \vdots & & \vdots \\ \lambda_1-\lambda_2 & \lambda_1-\lambda_2 & \cdots & r-\lambda_2 \end{bmatrix}.$$

From (5.0) and (5.1)

$$|D| = rk(r-\lambda_1)^{n-1}, \qquad |C| = (\lambda_2-\lambda_1)(r-\lambda_1)^{n-2}.$$

If from $M_2$ we strike out the $2n$th, $\cdots$, $mn$th rows and columns, then the determinant of the resulting matrix is

$$|D|\,|C|^{m-1} = rk(\lambda_2-\lambda_1)^{m-1}(r-\lambda_1)^{mn-2m+1} \neq 0,$$

since $\lambda_1 \neq \lambda_2$ and $r - \lambda_1 \neq 0$. Hence

$$\text{Rank } M_2 \geqq n + (m-1)(n-1) = v - m + 1. \tag{5.6}$$

But

$$\text{Rank } M_2 = \text{Rank } NN' \leqq b. \tag{5.7}$$

Hence the following theorem.

THEOREM 5. *For a semi-regular GD design* $b \geqq v - m + 1$.

If the design is resolvable, that is, the blocks can be divided into $r$ groups, of $b/r$ blocks each, such that each group of blocks gives a complete replication, then

$$\text{Rank } NN' = \text{Rank } M_2 \leqq b - r + 1, \tag{5.8}$$

since in $N$ the sum of the columns corresponding to a complete replication must give a column consisting of unities. Thus not more than $b - r + 1$ column vectors are independent. Hence the following corollary.

COROLLARY. *For a resolvable semi-regular GD design* $b \geqq v - m + r$.

**6. Regular GD designs.** For a regular GD design we have by definition

$$r > \lambda_1, \qquad rk - v\lambda_2 > 0. \tag{6.0}$$

From (3.7), $|NN'| > 0$. Hence

$$v = \text{Rank } NN' = \text{Rank } N \leqq b. \tag{6.1}$$

This gives

THEOREM 6. *For a regular GD design* $b \geqq v$.

If the design is resolvable then as before Rank $N \leqq b - r + 1$. Hence we can state the

COROLLARY. *For a resolvable regular GD design* $b \geqq v + r - 1$.

A GD design is said to be symmetrical if $b = v$, and in consequence $r = k$. Shrikhande [9] and Chowla and Ryser [10] have obtained conditions necessary for the existence of symmetrical balanced incomplete block designs. We shall extend their results to symmetrical regular GD designs. With this in view we give a brief resume of the important properties of the Hilbert norm residue symbol $(a, b)_p$, where $a$ and $b$ are rational numbers and $p$ is a prime, and of the Minkowski-Hasse invariant [11], [12] $C_p(A)$ of a symmetric matrix with rational elements. We shall use these symbols in the sense defined by Pall [13]. Detailed proofs and a more general theory are available in a book by Jones [14].

**7. The Hilbert norm residue symbol** $(a, b)_p$. If $a$ and $b$ are any nonzero rational numbers, we define the Hilbert symbol $(a, b)_p$ to have the value $+1$ or $-1$ according as the congruence

$$ax^2 + by^2 \equiv 1 (\text{mod } p^r) \tag{7.0}$$

has or has not for each value $r$, rational solutions $x_r$ and $y_r$. Here $p$ is any prime including the conventional prime $p_\infty = \infty$.

It is clear that the value of $(a, b)_p$ is unchanged if we replace $a$ and $b$ by their square-free parts, that is, if $s$ and $t$ are rational numbers then

$$(as^2, bt^2)_p = (a, b)_p. \tag{7.1}$$

Pall [13] has given the following formulae for calculating $(a, b)_p$ in general, where $(a \mid p)$ is the well known Legendre symbol, and $m$ and $m'$ are prime to $p$.

$$(a, b)_\infty = -1, \qquad \text{if and only if } a \text{ and } b \text{ are negative;} \tag{7.2}$$

$$(p^\alpha m, p^{\alpha'} m')_p = (-1 \mid p)^{\alpha\alpha'} (m \mid p)^{\alpha'} (m' \mid p)^\alpha \text{ if } p > 2; \tag{7.3}$$

$$(2^\alpha m, 2^{\alpha'} m')_2 = (2 \mid m)^{\alpha'} (2 \mid m')^\alpha (-1)^{(m-1)(m'-1)/4}. \tag{7.4}$$

The following properties of $(a, b)_p$ stated by Pall can be easily verified.

$$(a, b)_p = (b, a)_p; \tag{7.5}$$

$$(a, -a)_p = 1, \qquad (a, a)_p = (a, -1)_p; \tag{7.6}$$

$$(a, b_1 b_2)_p = (a, b_1)_p (a, b_2)_p, \qquad (a_1 a_2, b)_p = (a_1, b)_p (a_2, b)_p. \tag{7.7}$$

For any odd prime $p$, $(a, b)_p$ is evidently $+1$ unless $p$ actually appears in $a$ or $b$. For a given $a$, $b$, $(a, b)_p$ is $-1$ for only a finite even number (possibly zero) of primes $p$ counting in $p_\infty = \infty$. This is expressed by writing

$$\prod_p (a, b)_p = +1. \tag{7.8}$$

**8. The Hasse-Minkowski invariant $C_p(A)$.** Instead of considering invariants of quadratic forms $f = \sum a_{ij}x_ix_j$ we might consider the invariants to refer to the corresponding symmetric matrices $(a_{ij})$.

Let $A = (a_{ij})$ be any $n \times n$ matrix with rational elements. The matrix $B$ is said to be rationally congruent to $A$, written $A \sim B$, provided there exists a nonsingular matrix $C$ with rational elements such that

$$A = C'BC, \tag{8.0}$$

where $C'$ is the transpose of $C$. If $D_i$ $(i = 1, 2, \cdots, n)$ denotes the leading principal minor determinant of order $i$ in the matrix $A$, then if none of the $D_i$ vanishes, the quantity

$$C_p = C_p(A) = (-1, -D_n)_p \prod_{i=1}^{n-1} (D_i, -D_{i+1})_p \tag{8.1}$$

is invariant for all matrices rationally congruent to $A$ [11].

The invariant $C_p$ may be expressed in a more symmetrical form as follows:

$$\begin{aligned} C_p(A) &= (-1, -1)_p\,(D_n, -1)_p \prod_{i=1}^{n-1} (D_i, D_{i+1})_p (D_i, -1)_p \\ &= (-1, -1)_p (D_1, -1)_p \prod_{i=1}^{n-1} (D_{i+1}, D_i)_p (D_{i+1}, -1)_p \\ &= (-1, -1)_p \prod_{i=0}^{n-1} (D_{i+1}, -D_i)_p, \end{aligned} \tag{8.2}$$

where $D_0 = 1$.

If $A_{n-1}$ denotes the matrix obtained by leaving out the last row and column of $A$, we at once get from (8.2)

$$c_p(A) = C_p(A_{n-1})(D_n, -D_{n-1})_p. \tag{8.3}$$

From this and the properties of the Hilbert symbol we can at once deduce the following Lemma, stated by Pall [13] in equivalent form.

LEMMA I. *If $A$ and $B$ are symmetric matrices with rational elements, and*

$$U = \begin{bmatrix} A & \\ & B \end{bmatrix} \tag{8.35}$$

*is the direct sum of $A$ and $B$, then*

$$C_p(U) = (-1, -1)_p C_p(A) C_p(B)(|A|, |B|)_p. \tag{8.4}$$

LEMMA II. *If $\Delta_m$ is the direct sum of $A$ taken $m$ times, where $A$ is a symmetric matrix with rational elements,* that is,

$$\Delta_m = \begin{bmatrix} A & & & \\ & A & & \\ & & \ddots & \\ & & & A \end{bmatrix}, \tag{8.45}$$

*then*

$$C_p(\Delta_m) = \{ C_p(A) \}^m ( | A | , -1)_p^{m(m-1)/2}(-1, -1)_p^{m-1}. \tag{8.5}$$

PROOF. From Lemma I

$$\begin{aligned} C_p(\Delta_2) &= (-1, -1)_p \{ C_p(A) \}^2( | A | , | A | )_p \\ &= (-1, -1)_p \{ C_p(A) \}^2( | A | , -1)_p . \end{aligned}$$

Hence the lemma holds for $m = 2$. Supposing it to hold for $m - 1$, we have

$$\begin{aligned} C_p(\Delta_m) &= C_p(\Delta_{m-1})C_p(A)( | A |^{m-1}, | A | )_p(-1, -1)_p \\ &= C_p(\Delta_{m-1})C_p(A)( | A | , -1)_p^{m-1}(-1, -1)_p \\ &= \{ C_p(A) \}^m( | A | , -1)_p^{m(m-1)/2}(-1, -1)_p^{m-1}. \end{aligned} \tag{8.6}$$

COROLLARY. *If $d$ is a rational number and $\Delta_m$ is the diagonal matrix of order $m$ given by*

$$\Delta_m = \begin{bmatrix} d & & & \\ & d & & \\ & & \ddots & \\ & & & d \end{bmatrix}, \tag{8.65}$$

*then*

$$C_p(\Delta_m) = (-1, -1)_p(d, -1)_p^{m(m+1)/2}. \tag{8.7}$$

This follows at once from the main lemma by noting that

$$C_p(d) = (-1, -1)_p(d, -1)_p$$

when $d$ is regarded as a matrix of order one (cf. (8.2)).

We shall prove another lemma on a property of the Hilbert symbol which we shall use later.

LEMMA III. *For any two rational numbers* $a \neq 0$, $b \neq 0$, *and any prime* $p$,

$$(a, b)_p = (-ab, a + b)_p . \tag{8.8}$$

Let

$$A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \sim \begin{bmatrix} a + b & b \\ b & b \end{bmatrix}.$$

Hence from (8.2)

$$(-1, -1)_p(a, -1)_p(ab, -a)_p = (-1, -1)_p(a + b, -1)_p(ab, -a-b)_p ,$$

or

$$(a, -1)_p(b, a)_p(b, -1)_p = (a + b, -1)_p(ab, -1)_p(ab, a + b)_p ,$$

or

$$(a, -1)_p(a, b)_p(b, -1)_p = (a + b, -ab)_p(a, -1)_p(b\ , -1)_p ,$$

or

$$(a, b)_p = (-ab, a + b)_p .$$

Putting $a = n + 1$, $b = -1$ we obtain the following result given as a lemma by Bruck and Ryser [15].

COROLLARY.

$$(n + 1, -1)_p = (n + 1, n)_p . \tag{8.9}$$

**9. Necessary conditions for the existence of symmetrical regular GD designs.** Consider a symmetrical regular GD design with parameters $v, b, r, k, m, n, \lambda_1, \lambda_2$, where

$$v = b = mn, \qquad r = k, \tag{9.0}$$

$$(n - 1)\lambda_1 + n(m - 1)\lambda_2 = r(r - 1), \tag{9.1}$$

$$Q = r - \lambda_1 > 0, \qquad P = r^2 - v\lambda_2 > 0. \tag{9.2}$$

Since the incidence matrix $N$ given by (3.3) is now a square matrix it follows from (3.7) that

$$|NN'| = |N|^2 = r^2P^{m-1}Q^{m(n-1)}. \tag{9.3}$$

It follows that $P^{m-1}Q^{m(n-1)}$ must be a perfect square. Hence the theorem

THEOREM 7. *A necessary condition for the existence of a symmetrical regular GD design with parameters* $v, b, r, k, m, n, \lambda_1, \lambda_2$ *satisfying* (9.0), (9.1) *and* (9.2) *is that* $P^{m-1}Q^{m(n-1)}$ *is a perfect square.*

Let $X = NN'$ be the matrix given by (3.5), then $X$ is rationally congruent to the unit matrix $I$. Hence a necessary condition for the design to exist is that

$C_p(X) = C_p(NN') = C_p(I)$ for every prime $p$. In order to calculate $C_p(X)$ we shall start by obtaining $C_p(Y)$, where $Y$ is the $n \times n$ matrix given by

$$Y = \begin{bmatrix} r-\lambda_2 & \lambda_1-\lambda_2 & \cdots & \lambda_1-\lambda_2 \\ \lambda_1-\lambda_2 & r-\lambda_2 & \cdots & \lambda_1-\lambda_2 \\ \vdots & \vdots & & \vdots \\ \lambda_1-\lambda_2 & \lambda_1-\lambda_2 & \cdots & r-\lambda_2 \end{bmatrix}. \tag{9.4}$$

Referring to the derivation of (3.7) it is readily seen that $|Y| = PQ^{n-1}$. Let us set

$$Y_1 = \begin{bmatrix} Y & \\ & -\lambda \end{bmatrix},$$

where $\lambda = \lambda_1 - \lambda_2$.

Hence from (8.3),

$$\begin{aligned} C_p(Y_1) &= C_p(Y)(-\lambda PQ^{n-1}, -PQ^{n-1})_p \\ &= C_p(Y)(-\lambda, -1)_p(-\lambda, P)_p(-\lambda, Q)_p^{n-1}. \end{aligned} \tag{9.45}$$

If we add the last row in $Y_1$ to each of the other rows, and then the last column to each of the other columns, we get

$$Y_1 \sim Y_2 = \begin{bmatrix} r-\lambda_1 & 0 & \cdots & 0 & -\lambda \\ 0 & r-\lambda_1 & \cdots & 0 & -\lambda \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & & r-\lambda_1 & -\lambda \\ -\lambda & -\lambda & \cdots & -\lambda & -\lambda \end{bmatrix}.$$

Since $|Y_2| = |Y_1| = -\lambda PQ^{n-1}$ it follows from (8.3) and (8.7) that

$$\begin{aligned} C_p(Y_1) = C_p(Y_2) &= (-1,-1)_p(Q,-1)_p^{n(n+1)/2}(-\lambda PQ^{n-1}, -Q^n)_p \\ &= (-1,-1)_p(Q,-1)_p^{n(n+1)/2}(-\lambda,-1)_p(-\lambda,Q)_p^n \\ &\quad \cdot (P,-1)_p(P,Q)_p^n(Q,-1)_p^{n-1}. \end{aligned} \tag{9.5}$$

Comparing (9.45) and (9.5) we finally get

$$C_p(Y) = (-1,-1)_p(Q,-1)_p^{n(n-1)/2}(P,\lambda)_p(Q,\lambda)_p(P,Q)_p^n, \tag{9.6}$$

where $\lambda = \lambda_1 - \lambda_2$.

To calculate $C_p(X) = C_p(NN')$ we now proceed as follows. Set

$$X_1 = \begin{bmatrix} X & \\ & -\lambda_2 \end{bmatrix}.$$

From (8.3) and (9.3),

$$C_p(X_1) = C_p(X)(-\lambda_2 r^2 P^{m-1} Q^{m(n-1)}, -r^2 P^{m-1} Q^{m(n-1)})_p$$
$$= C_p(X)(-\lambda_2, -1)_p, \tag{9.65}$$

since from Theorem 7 $P^{m-1}Q^{m(n-1)}$ is a perfect square.

Again we may transform $X_1$ to the rationally congruent matrix $X_2$, by adding the last row of $X_1$ to each of the other rows, and then the last column of $X_1$ to each of the other columns. Thus

$$X_1 \sim X_2 = \begin{bmatrix} Y & & & & L \\ & Y & & & L \\ & & \ddots & & \vdots \\ & & & Y & L \\ L' & L' & \cdots & L' & -\lambda_2 \end{bmatrix},$$

where $Y$ is given by (9.4), $L$ is an $n \times 1$ column matrix with each element as $-\lambda_2$ and $L'$ is the transpose of $L$. We note that

$$|X_2| = |X_1| = -\lambda_2 r^2 P^{m-1} Q^{m(n-1)}.$$

Hence from (8.3) and (8.5),

$$C_p(X_1) = C_p(X_2) = \{C_p(Y)\}^m (PQ^{n-1}, -1)_p^{m(m-1)/2} (-1, -1)_p^{m-1}$$
$$\cdot (-\lambda_2 r^2 P^{m-1} Q^{m(n-1)}, -P^m Q^{m(n-1)})_p$$
$$= \{C_p(Y)\}^m (PQ^{n-1}, -1)_p^{m(m-1)/2} (-1, -1)_p^{m-1} (-\lambda_2, -1)_p (-\lambda_2, P)_p, \tag{9.7}$$

since $P^{m-1}Q^{m(n-1)}$ is a perfect square.

Comparing (9.65) and (9.7)

$$C_p(X) = C_p(NN') = \{C_p(Y)\}^m (PQ^{n-1}, -1)_p^{m(m-1)/2} (-1, -1)_p^{m-1} (-\lambda_2, P)_p.$$

Substituting from (9.6),

$$C_p(NN') = (-1, -1)_p (Q, -1)_p^{m(n-1)(m+n-1)/2} (P, -1)_p^{m(m-1)/2}$$
$$\cdot (PQ, \lambda_1 - \lambda_2)_p^m (P, Q)_p^{mn} (P, -\lambda_2)_p. \tag{9.8}$$

Now

$$(P, -1)_p^{m-1} (P, Q)_p^{m(n-1)} = (P, P^{m-1})_p (P, Q^{m(n-1)})_p$$
$$= (P, P^{m-1} Q^{m(n-1)})_p \tag{9.85}$$
$$= +1.$$

Also from (9.1)

$$n(\lambda_1 - \lambda_2) = P - Q.$$

Therefore,

$$(PQ, \lambda_1 - \lambda_2)_p = (PQ, P - Q)_p(PQ, n)_p$$
$$= (P, -Q)_p(P, n)_p(Q, n)_p$$

from (8.8), whence

$$(PQ, \lambda_1 - \lambda_2)_p = (P, -1)_p(P, Q)_p(P, n)_p(Q, n)_p .$$

Hence finally

$$C_p(NN') = (-1, -1)_p(Q, -1)_p^{m(n-1)(m+n-1)/2}(P, -1)_p^{m(m-1)/2} \cdot (P, n)_p^m(Q, n)_p^m(P, \lambda_2)_p . \tag{9.9}$$

We thus get

THEOREM 8. *A necessary condition for the existence of a symmetrical regular GD design with parameters $v, b, r, k, m, n, \lambda_1, \lambda_2$ satisfying* (9.0), (9.1) *and* (9.2) *is that $C_p(NN')$, given by* (9.9), *is equal to $C_p(I)$ for all primes $p$, where $I$ is the unit matrix.*

Note that from (7.2), (7.4), (7.6) and (8.2),

$$C_p(I) = (-1, -1)_p = -1 \quad \text{if } p = 2 \text{ or } p_\infty,$$
$$= +1 \quad \text{if } p \text{ is an odd prime.} \tag{9.35}$$

Confining ourselves only to the case when $p$ is an odd prime, we can combine Theorems 7 and 8 into the following single theorem.

THEOREM 9. *If a symmetrical regular GD design with parameters $v, b, r, k, m, n, \lambda_1, \lambda_2$ satisfying* (9.0), (9.1) *and* (9.2) *exists then*

(a) *if $m$ is even $P$ must be a perfect square and if further $m = 4t + 2$ and $n$ is even $(Q, -1)_p = +1$ for all odd primes $p$;*

(b) *if $m$ is odd and $n$ is even $Q$ is a perfect square, and $((-1)^\alpha n\lambda_2, P)_p = +1$ for all odd primes $p$, where $\alpha = m(m-1)/2$;*

(c) *if $m$ and $n$ are both odd $((-1)^\alpha n\lambda_2, P)_p((-1)^\beta n, Q)_p = +1$ for all odd primes $p$, where $\alpha = m(m-1)/2$ and $\beta = n(n-1)/2$.*

PROOF.

(a) If $m$ is even then in $P^{m-1}Q^{m(n-1)}$ the index of $Q$ is even and the index of $P$ is odd. Hence from Theorem 7, $P$ must be a perfect square. Noting that $(-1, -1)_p = +1$ for all odd primes $p$, $C_p(NN')$ reduces to

$$(Q, -1)_p^{m(n-1)(m+n-1)/2}.$$

If now $m$ is of the form $4t + 2$ and $n$ is even, then the index of $(Q, -1)_p$ is odd. Theorem 8 shows that $(Q, -1)_p = +1$.

(b) If $m$ is odd and $n$ is even then in $P^{m-1}Q^{m(n-1)}$ the index of $P$ is even and the index of $Q$ is odd. Hence from Theorem 7, $Q$ is a perfect square. Also $C_p(NN')$ reduces to $((-1)^\alpha n\lambda_2, P)_p$ for all odd primes $p$, if $\alpha = m(m-1)/2$. The result follows from Theorem 8.

(c) The result follows from Theorem 8 by noting that if $m$ and $n$ are both odd then for all odd primes $p$

$$\begin{aligned} C_p(NN') &= (n\lambda_2, P)_p(n, Q)_p && \text{if } m = 4t+1, \quad n = 4t+1; \\ &= (-n\lambda_2, P)_p(n, Q)_p && \text{if } m = 4t+3, \quad n = 4t+1; \\ &= (n\lambda_2, P)_p(-n, Q)_p && \text{if } m = 4t+1, \quad n = 4t+3; \\ &= (-n\lambda_2, P)_p(-n, Q)_p && \text{if } m = 4t+3, \quad n = 4t+3. \end{aligned}$$

We give below a table of some symmetrical regular GD designs whose impossibility can be proved by using Theorem 9. The last column gives the reason for impossibility. Thus for the design "Ref. no. 1," the statement $P = 5$, Th. 9(a) in the last column means that $P = 5$ contradicts Theorem 9(a). Hence the design is impossible.

*Some impossible symmetrical regular GD designs*

| Ref. no. | $v=b$ | $r=k$ | $m$ | $n$ | $\lambda_1$ | $\lambda_2$ | Reason for impossibility |
|---|---|---|---|---|---|---|---|
| 1 | 44 | 7 | 22 | 2 | 0 | 1 | $P = 5$, Th. 9(a) |
| 2 | 92 | 10 | 46 | 2 | 0 | 1 | $P = 8$, Th. 9(a) |
| 3 | 56 | 8 | 28 | 2 | 2 | 1 | $P = 8$, Th. 9(a) |
| 4 | 88 | 10 | 22 | 4 | 2 | 1 | $P = 12$, Th. 9(a) |
| 5 | 20 | 5 | 10 | 2 | 2 | 1 | $P = 5$, Th. 9(a) |
| 6 | 18 | 7 | 2 | 9 | 3 | 2 | $P = 13$, Th. 9(a) |
| 7 | 30 | 9 | 2 | 15 | 3 | 2 | $P = 21$, Th. 9(a) |
| 8 | 22 | 5 | 11 | 2 | 0 | 1 | $Q = 5$, Th. 9(b) |
| 9 | 42 | 7 | 21 | 2 | 2 | 1 | $Q = 5$, Th. 9(b) |
| 10 | 90 | 10 | 45 | 2 | 2 | 1 | $Q = 8$, Th. 9(b) |
| 11 | 30 | 8 | 15 | 2 | 0 | 2 | $Q = 8$, Th. 9(b) |
| 12 | 33 | 6 | 11 | 3 | 0 | 1 | $(-n\lambda_2, P)_3(-n, Q)_3 = -1$, Th. 9(c) |
| 13 | 93 | 10 | 31 | 3 | 0 | 1 | $(-n\lambda_2, P)_5(-n, Q)_5 = -1$, Th. 9(c) |
| 14 | 95 | 10 | 19 | 5 | 0 | 1 | $(-n\lambda_2, P)_5(n, Q)_5 = -1$, Th. 9(c) |
| 15 | 39 | 7 | 13 | 3 | 3 | 1 | $(n\lambda_2, P)_5(-n, Q)_5 = -1$, Th. 9(c) |
| 16 | 27 | 7 | 3 | 9 | 3 | 1 | $(-n\lambda_2, P)_{11}(n, Q)_{11} = -1$, Th. 9(c) |
| 17 | 69 | 9 | 23 | 3 | 3 | 1 | $(-n\lambda_2, P)_3(-n, Q)_3 = -1$, Th. 9(c) |
| 18 | 65 | 9 | 13 | 5 | 3 | 1 | $(n\lambda_2, P)_5(n, Q)_5 = -1$, Th. 9(c) |
| 19 | 15 | 6 | 5 | 3 | 3 | 2 | $(n\lambda_2, P)_3(-n, Q)_3 = -1$, Th. 9(c) |
| 20 | 35 | 9 | 7 | 5 | 3 | 2 | $(-n\lambda_2, P)_3(n, Q)_3 = -1$, Th. 9(c) |
| 21 | 33 | 10 | 11 | 3 | 0 | 3 | $(-n\lambda_2, P)_5(-n, Q)_5 = -1$, Th. 9(c) |

## REFERENCES

[1] R. C. Bose and K. R. Nair, "Partially balanced incomplete block designs," *Sankhyā*, Vol. 4 (1939), pp. 337–372.

[2] K. R. Nair and C. R. Rao, "A note on partially balanced incomplete bolck designs," *Science and Culture*, Vol. 7 (1942), pp. 568–569.

[3] K. R. Nair and C. R. Rao, "Incomplete block designs for experiments involving several groups of varieties," *Science and Culture*, Vol. 7 (1942), p. 625.

[4] C. R. Rao, "General method of analysis for incomplete block designs," *Jour. Amer. Stat. Assn.*, Vol. 42 (1947), pp. 541–561.
[5] F. Yates, "Incomplete randomized blocks," *Annals of Eugenics*, Vol. 7 (1936), pp. 121–140.
[6] R. C. Bose, "On the construction of balanced incomplete block designs," *Annals of Eugenics*, Vol. 9 (1939), pp. 353–399.
[7] R. A. Fisher, "An examination of the different possible solutions of a problem in incomplete blocks," *Annals of Eugenics*, Vol. 10 (1940), pp. 52–75.
[8] R. C. Bose, "A note on Fisher's inequality for balanced incomplete block designs," *Annals of Math. Stat.*, Vol. 20 (1949), pp. 619–620.
[9] S. S. Shrikhande, "The impossibility of certain symmetrical balanced incomplete block designs," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 106–111.
[10] S. Chowla and H. J. Ryser, "Combinatorial problems," *Canadian Jour. Math.*, Vol. 2 (1950), pp. 93–99.
[11] H. Hasse, "Über die Äquivalenz quadratischer Formen im Körper der rationalen Zahlen," *Jour. reine angew. Math.*, Vol. 152 (1923), pp. 205–224.
[12] H. Minkowski, *Gesammelte Abhandlungen*, Vol. 1, B. G. Teubner, Berlin, 1932, pp. 161–173.
[13] G. Pall, "The arithmetical invariants of quadratic forms," *Bull. Am. Math. Soc.*, Vol. 51 (1945), pp. 185–197.
[14] B. W. Jones, *The Arithmetic Theory of Quadratic Forms*, John Wiley and Sons, 1950.
[15] R. H. Bruck and H. J. Ryser, "The nonexistence of certain finite projective planes," *Canadian Jour. Math.*, Vol. 1 (1949), pp. 88–93.

# THE OPERATING CHARACTERISTIC OF THE CONTROL CHART FOR SAMPLE MEANS[1]

By Edgar P. King

*Carnegie Institute of Technology*[2]

**1. Summary.** In this paper we derive the operating characteristic of the control chart for sample means when process standards are unspecified. Under the null hypothesis the distribution of the process is $N(\mu, \sigma^2)$, where $\mu$ and $\sigma$ are fixed but unknown. Under the alternative the process mean is a random variable with a $N(\mu, \theta^2\sigma^2)$ distribution. Exact results are obtained for cases ranging from two samples of size 2 to four samples of 10. Bounds on the operating characteristic are obtained in particular cases ranging from five samples of 5 to 25 samples of 10.

**2. Introduction.** The usual procedure in constructing control charts from past data consists of the following steps [1].

(a) Classify the total number, $N$ say, of observations to be drawn from the process into $m$ samples of size $n$ according to some rational method of subgrouping.

(b) For each sample, calculate a mean and a range, plotting these values on separate charts in the order drawn.

(c) Using prescribed formulae based on the data collected and the sample size, calculate upper and lower control limits for each chart.

(d) If all the plotted points fall within the control limits on both charts, accept the hypothesis that the process is in a state of statistical control. Otherwise, reject this hypothesis.

In what follows we shall confine our attention to the control chart for means. Letting $x_{ij}$ denote the $j$th observation in the $i$th random sample of size $n$, we will be concerned with the following linear model

$$x_{ij} = \mu_i + \epsilon_{ij}, \qquad i = 1, 2, \cdots, m; j = 1, 2, \cdots, n,$$

where

(1) the $\epsilon_{ij}$ are statistically independent and distributed according to $N(0, \sigma^2)$;

(2) the $\mu_i$ are statistically independent and distributed according to $N(\mu, \theta^2\sigma^2)$;

(3) the $\mu_i$ are statistically independent of the $\epsilon_{ij}$;

(4) $\mu$ and $\sigma$ are fixed but unknown.

Let $\bar{x}_i$ and $r_i$ denote the mean and range, respectively, of the $i$th sample $x_{i1}, \cdots, x_{in}$. The usual control chart procedure prescribes that the limits on the $\bar{x}$-chart be set at $\bar{\bar{x}} \pm A_2\bar{r}$, where

$$\bar{\bar{x}} = \frac{1}{mn}\sum_{j=1}^{n}\sum_{i=1}^{m} x_{ij}, \qquad \bar{r} = \frac{1}{m}\sum_{i=1}^{m} r_i,$$

[1] Submitted in partial fulfillment of the requirements for the D.Sc. degree at Carnegie Institute of Technology.

[2] Now at Statistical Engineering Laboratory, National Bureau of Standards, Washington, D. C.

and $A_2$ is defined by $E(A_2\bar{r}) = 3\sigma/\sqrt{n}$ when the underlying distribution is normal. What we seek then is the probability that the inequalities

$$\bar{x} - A_2\bar{r} < \bar{x}_i < \bar{x} + A_2\bar{r}, \qquad i = 1, 2, \cdots, m,$$

be satisfied simultaneously. Defining

$$\beta(k) = P\left(-\frac{k}{3}A_2\bar{r} < \bar{x}_1 - \bar{x} < \frac{k}{3}A_2\bar{r}, \cdots, -\frac{k}{3}A_2\bar{r} < \bar{x}_m - \bar{x} < \frac{k}{3}A_2\bar{r}\right) \tag{2.1}$$

we wish to evaluate $\beta(3)$ under the hypotheses

$$H_0 : \theta = 0;$$

$$H_1 : \theta > 0.$$

The null hypothesis, $H_0$, is the hypothesis of statistical control. Our choice of the alternative $\theta > 0$ is motivated by practical considerations. Since shifts in the mean of an industrial process might occur at any time and be of any magnitude, an alternative which ordered the $\mu_i$ in a particular way would be of limited interest. By treating the $\mu_i$ as random variables we are taking into account the "average effect" of $m$ independent assignable causes of varying magnitude. Although we are unable to specify the size of any particular shift, a measure of the size of the $\mu_i$ as a group is given by the parameter $\theta$.

**3. Taylor's series expansion of $\beta(3)$.** The general method employed in calculating $\beta(3)$ is first to obtain

$$\begin{aligned}\beta_0(k) &= P\left(-k\frac{\sigma}{\sqrt{n}} < \bar{x}_1 - \bar{\bar{x}} < k\frac{\sigma}{\sqrt{n}}, \cdots, -k\frac{\sigma}{\sqrt{n}} < \bar{x}_m - \bar{\bar{x}} < k\frac{\sigma}{\sqrt{n}}\right) \\ &= P\left(-k < \frac{\bar{x}_1 - \bar{\bar{x}}}{\sigma/\sqrt{n}} < k, \cdots, -k < \frac{\bar{x}_m - \bar{\bar{x}}}{\sigma/\sqrt{n}} < k\right)\end{aligned} \tag{3.1}$$

as a function of $k$ for fixed $m$. We note that $\beta_0(k)$ is of considerable interest in its own right, since $\beta_0(3)$ represents the probability that the $\bar{x}$-chart will show control when $\sigma$ is known.

To evaluate $\beta(3)$ we multiply the conditional probability that (for fixed $\bar{r}$) $-A_2\bar{r} < \bar{x}_i - \bar{\bar{x}} < A_2\bar{r}$ $(i = 1, 2, \cdots, m)$ by the pdf of $\bar{r}$ and integrate on $\bar{r}$ from 0 to $\infty$. For fixed $\bar{r}$ we have

$$P(-A_2\bar{r} < \bar{x}_1 - \bar{\bar{x}} < A_2\bar{r}, \cdots)$$

$$= P\left(-\frac{A_2\bar{r}}{\sigma/\sqrt{n}} < \frac{\bar{x}_1 - \bar{\bar{x}}}{\sigma/\sqrt{n}} < \frac{A_2\bar{r}}{\sigma/\sqrt{n}}, \cdots\right) = \beta_0\left(\frac{\sqrt{n}\,A_2\bar{r}}{\sigma}\right).$$

Let $g_{mn}(\bar{r};\ \sigma)$ denote the pdf of the average range of $m$ random samples of $n$ from a $N(\mu,\ \sigma^2)$ universe. Then

$$\beta(3) = \int_0^\infty g_{mn}(w;\ \sigma)\beta_0\left(\sqrt{n}\,A_2\frac{w}{\sigma}\right)dw.$$

Letting $g_{mn}(w; 1) = g_{mn}(w)$, we have

$$g_{mn}(w; \sigma) = \frac{1}{\sigma} g_{mn}\left(\frac{w}{\sigma}\right).$$

Thus

$$\beta(3) = \int_0^\infty \frac{1}{\sigma} g_{mn}\left(\frac{w}{\sigma}\right) \beta_0\left(\sqrt{n}\, A_2 \frac{w}{\sigma}\right) dw.$$

Finally, letting $\bar{r} = w/\sigma$ yields

$$\beta(3) = \int_0^\infty g_{mn}(\bar{r}) \beta_0(\sqrt{n} A_2 \bar{r})\, d\bar{r}. \tag{3.2}$$

We now expand $\beta_0(\sqrt{n}\ A_2\bar{r})$ in a Taylor's series with remainder about the point $\bar{r} = d_2$, where $d_2\sigma$ denotes the expected value of the range in a random sample of $n$ from a $N(\mu, \sigma^2)$ population. Clearly

$$\beta_0(\sqrt{n}A_2\bar{r}) = \beta_0(3) + b_1(\bar{r} - d_2) + \cdots + \frac{b_p}{p!}(\bar{r} - d_2)^p + R'(\xi), \tag{3.3}$$

where the $b_i$ are the $i$th derivatives of $\beta_0(\sqrt{n}\ A_2\bar{r})$ with respect to $\bar{r}$, evaluated at $\bar{r} = d_2$, and

$$R'(\xi) = \frac{b_{p+1}(\xi)}{(p+1)!}(\bar{r} - d_2)^{p+1}, \qquad \xi = d_2 + \alpha(\bar{r} - d_2), \qquad 0 \leqq \alpha \leqq 1. \tag{3.4}$$

This expansion is valid since all the $b_i$ are continuous. Now, taking the expectation of both sides of (3.3) yields

$$\int_0^\infty g_{mn}(\bar{r})\beta_0(\sqrt{n}A_2\bar{r})\, d\bar{r} = \beta_0(3) \int_0^\infty g_{mn}(\bar{r})\, d\bar{r} + b_1 \int_0^\infty (\bar{r} - d_2) g_{mn}(\bar{r})\, d\bar{r} + \cdots + \int_0^\infty R'(\xi) g_{mn}(\bar{r})\, d\bar{r}$$

But since the first integral on the right side is unity and the second zero, we have

$$\beta(3) = \beta_0(3) + \frac{b_2}{2!}\mu_2(\bar{r}) + \cdots + \frac{b_p}{p!}\mu_p(\bar{r}) + R, \tag{3.5}$$

where the $\mu_i(\bar{r})$ are central moments of average range, and

$$R = \int_0^\infty \frac{b_{p+1}(\xi)}{(p+1)!}(\bar{r} - d_2)^{p+1} g_{mn}(\bar{r})\, d\bar{r}. \tag{3.6}$$

For the actual numerical computation of $\beta(3)$, relation (3.5) is used in the form

$$\beta(3) = \beta_0(3) + \frac{b_2}{2!}\frac{\mu_2(r)}{m} + \frac{b_3}{3!}\frac{\mu_3(r)}{m^2} + \frac{b_4}{4!}\frac{1}{m^3}[\mu_4(r) + 3(m-1)\mu_2^2(r)] + \cdots, \tag{3.7}$$

where the $\mu_i(r)$ are the $i$th central moments of the range in samples of $n$ from a $N(\mu, 1)$ population. In each case, a sufficient number of terms of (3.7) is used to insure that the remainder, $R$, is small.

**4. Derivation of $\beta_0(k)$.** As indicated in the previous section, our first task is to obtain an expression for $\beta_0(k)$. It is apparent that it can be written as

$$\beta_0(k) = P\left(\frac{\bar{x}_{(m)} - \bar{\bar{x}}}{\sigma/\sqrt{n}} < k, \frac{\bar{\bar{x}} - \bar{x}_{(1)}}{\sigma/\sqrt{n}} < k\right), \tag{4.1}$$

where $\bar{x}_{(m)}$ and $\bar{x}_{(1)}$ are the largest and smallest sample means, respectively. To obtain a tractable expression for this probability we shall first investigate the joint distribution of $\frac{\bar{x}_{(m)} - \bar{\bar{x}}}{\sigma/\sqrt{n}}$ and $\frac{\bar{\bar{x}} - \bar{x}_{(1)}}{\sigma/\sqrt{n}}$. For this purpose let us consider the joint pdf of the ordered variates $\bar{x}_{(1)} < \bar{x}_{(2)} < \cdots < \bar{x}_{(m)}$. Since $\bar{x}_i = \mu_i + \bar{\epsilon}_i$, where $\bar{\epsilon}_i = \sum_{j=1}^{n} \epsilon_{ij}/n$, we see that the distribution of $\bar{x}_i$ is normal with mean $\mu$ and variance $(\theta^2 + 1/n)\sigma^2$. Assuming $\mu = 0$ (with no loss of generality) we have

$$f(\bar{x}_{(1)}, \cdots, \bar{x}_{(m)}) = \frac{m!}{(2\pi)^{m/2}\left(\frac{1 + n\theta^2}{n}\right)^{m/2}\sigma^m} \exp\left[-\frac{n}{2(1 + n\theta^2)\sigma^2}\sum_{i=1}^{m}\bar{x}_{(i)}^2\right]$$

as the joint pdf of $\bar{x}_{(1)} < \bar{x}_{(2)} < \cdots < \bar{x}_{(m)}$. To obtain the joint distribution of $\frac{\bar{x}_{(m)} - \bar{\bar{x}}}{\sigma/\sqrt{n}}$ and $\frac{\bar{\bar{x}} - \bar{x}_{(1)}}{\sigma/\sqrt{n}}$, let

$$\frac{\sigma}{\sqrt{n}}\sqrt{2}\, v_1 = -\bar{x}_{(2)} + \bar{x}_{(3)},$$

$$\frac{\sigma}{\sqrt{n}}\sqrt{3\cdot 2}\, v_2 = -\bar{x}_{(2)} - \bar{x}_{(3)} + 2\bar{x}_{(4)},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{\sigma}{\sqrt{n}}\sqrt{(m-3)(m-2)}\, v_{m-3} = -\bar{x}_{(2)} - \bar{x}_{(3)} - \cdots - \bar{x}_{(m-2)} + (m-3)\,\bar{x}_{(m-1)},$$

$$\frac{\sigma}{\sqrt{n}}\sqrt{(m-1)(m-2)}\, v_{m-2} = -(m-2)\bar{x}_{(1)} + \bar{x}_{(2)} + \cdots + \bar{x}_{(m-1)},$$

$$\frac{\sigma}{\sqrt{n}}\sqrt{m(m-1)}\, v_{m-1} = -\bar{x}_{(1)} - \bar{x}_{(2)} - \cdots - \bar{x}_{(m-1)} + (m-1)\,\bar{x}_{(m)},$$

$$\frac{\sigma}{\sqrt{n}}\sqrt{m}\, v_m = \bar{x}_{(1)} + \bar{x}_{(2)} + \cdots + \bar{x}_{(m-1)} + \bar{x}_{(m)}.$$

This transformation gives

$$g(v_1, \cdots, v_m) = \frac{m!}{(2\pi)^{m/2}(1 + n\theta^2)^{m/2}} \exp\left[-\frac{1}{2(1 + n\theta^2)}\sum_{i=1}^{m} v_i^2\right]$$

as the joint pdf of the $v_i$, defined over the region $S$ given by

$$S\begin{cases} v_1 > 0, \quad v_i < \sqrt{\dfrac{i+2}{i}}\, v_{i+1}, & (i = 1, \cdots, m-4), \\ \sqrt{m(m-2)}\, v_{m-1} > \sqrt{(m-1)(m-3)}\, v_{m-3} + v_{m-2}, \\ \sqrt{\dfrac{m-1}{m-2}}\, v_{m-2} > \dfrac{v_1}{\sqrt{2}} + \dfrac{v_2}{\sqrt{3\cdot 2}} + \cdots + \dfrac{v_{m-3}}{\sqrt{(m-2)(m-3)}}. \end{cases}$$

Now, we integrate out $v_m$ and let

$$v_i = u_i, \qquad (i = 1, \cdots, m-3),$$

$$v_{m-2} = -\frac{u}{\sqrt{(m-1)(m-2)}} + \sqrt{\frac{m-1}{m-2}}\, v,$$

$$v_{m-1} = \sqrt{\frac{m}{m-1}}\, u,$$

where in terms of the original variables

$$\begin{cases} u = \dfrac{\bar{x}_{(m)} - \bar{x}}{\sigma/\sqrt{n}}, \\ v = \dfrac{\bar{x} - \bar{x}_{(1)}}{\sigma/\sqrt{n}}. \end{cases} \tag{4.2}$$

The Jacobian of this transformation is $|J| = \sqrt{m/(m-2)}$. The joint density of $u_1, u_2, \cdots, u_{m-3}, u, v$ is finally obtained as

$$h_m(u_1, \cdots, u_{m-3}, u, v) = \frac{m!\sqrt{\dfrac{m}{m-2}}}{(2\pi)^{(m-1)/2}(1+n\theta^2)^{(m-1)/2}} \cdot \exp\left[-\frac{1}{2(1+n\theta^2)}\sum_{i=1}^{m-3} u_i^2\right] \exp\left[-\frac{\Lambda_m}{2(1+n\theta^2)}\right] \tag{4.3}$$

over $S'$, where

$$\Lambda_m = \frac{m-1}{m-2}(u^2+v^2) - \frac{2}{m-2}uv,$$

$$S' \text{ is } \begin{cases} u_1 > 0, u > 0, v > 0, u_i < \sqrt{\dfrac{i+2}{i}}\, u_{i+1}, & (i = 1, \cdots, m-4), \\ \dfrac{u_1}{\sqrt{2}} + \dfrac{u_2}{\sqrt{3\cdot 2}} + \cdots + \dfrac{u_{m-3}}{\sqrt{(m-2)(m-3)}} < \dfrac{1}{m-2}[(m-1)v - u], \\ \sqrt{(m-2)(m-3)}\, u_{m-3} < (m-1)u - v. \end{cases}$$

For the cases $m = 2$ through $m = 4$, $\beta_0(k)$ is obtained by integrating (4.3) over the range $u_1, \cdots, u_{m-3} \in S'$, $u < k$, $v < k$. $\beta(3)$ is then evaluated from the series (3.7). For cases where $m > 4$, bounds on the desired probabilities are obtained by methods described in Section 5*D*.

**5. Exact evaluation of $\beta_0(k)$ and $\beta(3)$.**

*A. Case* $m = 2$. In this case, $\beta_0(k)$ becomes simply

$$\begin{aligned} \beta_0(k) &= P\left(-k < \frac{\bar{x}_1 - \bar{x}_2}{2\sigma/\sqrt{n}} < k\right) \\ &= \Phi\left(\frac{\sqrt{2}k}{\sqrt{1+n\theta^2}}\right) - \Phi\left(-\frac{\sqrt{2}k}{\sqrt{1+n\theta^2}}\right), \end{aligned} \tag{5.1}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. The series expansion for $\beta(3)$ becomes

$$\beta(3) = \beta_0(3) + \frac{b_2}{4}\mu_2(r) + \frac{b_3}{24}\mu_3(r) + \frac{b_4}{192}(\mu_4(r) + 3\mu_2^2(r)) + \cdots, \tag{5.2}$$

where

$$b_i = \left[\frac{d_i}{d\bar{r}^i}\beta_0(\sqrt{n}\, A_2\bar{r})\right]_{\bar{r}=d_2}.$$

The moments $\mu_i(r)$ are obtained from Hartley and Pearson [2]. Expression (5.2) is used to evaluate $\beta(3)$ for $n = 5$ and $n = 10$. For the sake of completeness the case $n = 2$ is also included. In this case, the pdf of the range is

$$h(r) = \frac{1}{\sqrt{\pi}\sigma} e^{-(r^2/(4\sigma^2))} \qquad (r > 0).$$

Hence, by the standard convolution (see [4], p. 191),

$$g_{22}(\bar{r};\sigma) = 2\int_0^{2\bar{r}} h(2\bar{r} - t)\, h(t)\, dt,$$

which reduces to

$$g_{22}(\bar{r};\sigma) = \frac{4}{\sqrt{2\pi}\sigma} e^{-(\bar{r}^2/(2\sigma^2))}\left[\Phi\left(\frac{\bar{r}}{\sigma}\right) - \Phi\left(-\frac{\bar{r}}{\sigma}\right)\right].$$

Hence

$$\begin{aligned}\beta(3) &= \int_0^{\infty} g_{22}(\bar{r})\beta_0(\sqrt{2}A_2\bar{r})\, d\bar{r} \\ &= \int_0^{\infty} \frac{4}{\sqrt{2\pi}} e^{-\frac{1}{2}\bar{r}^2}[\Phi(\bar{r}) - \Phi(-\bar{r})] \\ &\qquad \cdot \left[\Phi\left(\frac{2A_2\bar{r}}{\sqrt{1 + 2\theta^2}}\right) - \Phi\left(-\frac{2A_2\bar{r}}{\sqrt{1 + 2\theta^2}}\right)\right] d\bar{r}.\end{aligned}$$

$1 - \beta(3)$ is then evaluated by numerical integration.

The results for $m = 2$ are summarized in Tables I and II below.

*B. Case $m = 3$.* We find from (4.3) that the joint density of $u$ and $v$ is given by

$$f_3(u, v) = \frac{3!\sqrt{3}}{2\pi(1 + n\theta^2)} \exp\left[-\frac{1}{2(1 + n\theta^2)}(u^2 - uv + v^2)\right] \text{ over } S', \tag{5.3}$$

where

$$S' \text{ is } \begin{cases} 0 < v < 2u, \\ 0 < u < 2v. \end{cases}$$

Then

$$\beta_0(k) = 2\int_0^{k} du \int_{\frac{1}{2}u}^{u} f_3(u, v)dv,$$

by virtue of the fact that $f_2(u, v)$ is symmetric about the line $u = v$. This integral reduces to

$$\beta_0(k) = 6\sqrt{\frac{3}{\pi}}\int_0^{k/(2\sqrt{1+n\theta^2})} e^{-3t^2}F_2(t)\,dt, \tag{5.4}$$

where $F_2(x)$ is defined by

$$F_2(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt.$$

$F_n(x)$ is the cdf of the extreme deviation from the sample mean in samples of $n$ from a $N(0, 1)$ population. This function has been tabulated by Grubbs [3]. $\beta_0(3)$

TABLE I

$\beta_0(3)$ *for* $m = 2$

| $\theta$ | $n = 2$ | $n = 5$ | $n = 10$ |
|---|---|---|---|
| 0 | 1.00 | 1.00 | 1.00 |
| .5 | 1.00 | 1.00 | .98 |
| 1.0 | .99 | .92 | .80 |
| 1.5 | .93 | .77 | .62 |
| 2.0 | .84 | .64 | .49 |
| 2.5 | .75 | .55 | .40 |
| 3.0 | .67 | .47 | .34 |

TABLE II

$\beta(3)$ *for* $m = 2$

| $\theta$ | $n = 2$ | $n = 5$ | $n = 10$ |
|---|---|---|---|
| 0 | .96 | 1.00 | 1.00 |
| .5 | .94 | .98 | .96 |
| 1.0 | .89 | .89 | .79 |
| 1.5 | .82 | .75 | .61 |
| 2.0 | .75 | .63 | .49 |
| 2.5 | .67 | .54 | .40 |
| 3.0 | .61 | .46 | .34 |

is evaluated from (5.4) by numerical integration and $\beta(3)$ obtained from the series (3.7). The results for this case are summarized in Tables III and IV.

*C. Case* $m = 4$. From (4.3) we obtain the joint pdf of $u_1$, $u$, and $v$ as

$$h_4(u_1, u, v) = \frac{4!\sqrt{\frac{4}{2}}}{(2\pi)^{\frac{1}{2}}(1 + n\theta^2)^{3/2}} e^{-[u_1^2/(2(1+n\theta^2))]}e^{-[\Lambda_4/(2(1+n\theta^2))]} \quad \text{over} \quad S, \tag{5.6}$$

where

$$\Lambda_4 = 3\tfrac{1}{2}(u^2 + v^2) - uv,$$

$$S \text{ is } \begin{cases} u_1 > 0,\ u > 0,\ v > 0, \\ \sqrt{2}\,u_1 < 3v - u, \\ \sqrt{2}\,u_1 < 3u - v. \end{cases}$$

To obtain $f_4(u, v)$, the joint density of $u$ and $v$, we integrate over the range of $u_1$, with the following result

$$f_4(u, v) = \begin{cases} C_1 e^{-[\Lambda_4/(2(1+n\theta^2))]}\displaystyle\int_0^{(3v-u)/\sqrt{2}} e^{-[u_1^2/(2(1+n\theta^2))]}\,du, & \text{for } v < u < 3v, \\ C_1 e^{-[\Lambda_4/(2(1+n\theta^2))]}\displaystyle\int_0^{(3u-v)/\sqrt{2}} e^{-[u_1^2/(2(1+n\theta^2))]}\,du, & \text{for } u < v < 3u, \end{cases}$$

where

$$C_1 = \frac{4!\sqrt{\frac{4}{2}}}{[2\pi(1+n\theta^2)]^{\frac{3}{2}}}.$$

But

$$\int_0^{(3v-u)/\sqrt{2}} e^{-u_1^2/(2(1+n\theta^2))}\,du_1 = \sqrt{2}\sqrt{1+n\theta^2}\int_0^{(3v-u)/(2\sqrt{1+n\theta^2})} e^{-t^2}\,dt$$

$$= \sqrt{\frac{\pi}{2}}\sqrt{1+n\theta^2}\,F_2\left(\frac{3v-u}{2\sqrt{1+n\theta^2}}\right).$$

TABLE III

$\beta_0(3)$ *for* $m = 3$

| $\theta$ | $n = 2$ | $n = 5$ | $n = 10$ |
|---|---|---|---|
| 0 | 1.00 | 1.00 | 1.00 |
| .5 | .99 | .96 | .88 |
| 1.0 | .92 | .71 | .48 |
| 1.5 | .74 | .45 | .27 |
| 2.0 | .56 | .30 | .17 |
| 2.5 | .42 | .21 | .11 |
| 3.0 | .32 | .15 | .08 |

TABLE IV

$\beta(3)$ *for* $m = 3$

| $\theta$ | $n = 5$ | $n = 10$ |
|---|---|---|
| 0 | .99 | 1.00 |
| .5 | .94 | .86 |
| 1.0 | .69 | .48 |
| 1.5 | .45 | .27 |
| 2.0 | .30 | .17 |
| 2.5 | .21 | .11 |
| 3.0 | .15 | .08 |

In exactly the same manner

$$\int_0^{(3u-v)/\sqrt{2}} e^{-[u_1^2/(2(1+n\theta^2))]}\,du_1 = \sqrt{\frac{\pi}{2}}\sqrt{1+n\theta^2}\,F_2\left(\frac{3u-v}{2\sqrt{1+n\theta^2}}\right).$$

Therefore

$$f_4(u,v) = C_1\sqrt{\frac{\pi}{2}}\sqrt{1+n\theta^2}\,e^{-[A_4/(2(1+n\theta^2))]}\,G(u,v),$$

where

$$G(u,v) = \begin{cases} F_2\left(\dfrac{3u-v}{2\sqrt{1+n\theta^2}}\right) & \text{for} \quad u < v < 3u, \\ F_2\left(\dfrac{3v-u}{2\sqrt{1+n\theta^2}}\right) & \text{for} \quad v < u < 3v. \end{cases}$$

Then

$$\beta_0(k) = 2\int_0^k du\int_{u/3}^u f_4(u,v)\,dv,$$

since $f_4(u, v)$ is symmetric in $u$ and $v$. After some reduction we find

$$\beta_0(k) = 12\frac{\sqrt{\frac{4}{3}}}{\sqrt{2\pi}}\int_0^{2k/(3\sqrt{1+n\theta^2})} e^{-(3t^2)/2} F_3(t)\,dt, \tag{5.7}$$

where

$$F_3(t) = \frac{3\sqrt{3}}{2\sqrt{\pi}}\int_0^t e^{-(3x^2/4)} F_2\left(\frac{3x}{2}\right) dx.$$

As before, $\beta_0(k)$ is evaluated by integrating (5.7) numerically and $\beta(3)$ is obtained from the series expansion. The results for this case are summarized in Tables V and VI.

TABLE V

$\beta_0(3)$ *for* $m = 4$

| $\theta$ | $n = 2$ | $n = 5$ | $n = 10$ |
|---|---|---|---|
| 0 | 1.00 | 1.00 | 1.00 |
| .5 | .98 | .91 | .76 |
| 1.0 | .82 | .52 | .27 |
| 1.5 | .56 | .25 | .11 |
| 2.0 | .34 | .12 | .05 |
| 2.5 | .22 | .07 | .03 |
| 3.0 | .14 | .05 | .02 |

TABLE VI

$\beta(3)$ *for* $m = 4$

| $\theta$ | $n = 5$ | $n = 10$ |
|---|---|---|
| 0 | .99 | 1.00 |
| .5 | .87 | .74 |
| 1.0 | .51 | .27 |
| 1.5 | .26 | .11 |
| 2.0 | .13 | .05 |
| 2.5 | .07 | .03 |
| 3.0 | .05 | .02 |

*D. Case* $m > 4$. In this section we derive upper and lower bounds for $\beta_0(k)$ and $\beta(3)$. Recalling that $\beta_0(k) = P(u < k, v < k)$ we immediately obtain the upper bound $\beta_0(k) \leqq P(u < k)$. But since $u = (\bar{x}_{(m)} - \bar{\bar{x}})/(\sigma/\sqrt{n})$, we see that $u/\sqrt{1 + n\theta^2}$ is distributed as the extreme deviation from the sample mean in samples of $m$ from a $N(0, 1)$ population. Hence, using Grubbs' notation [3]

$$\beta_0(k) \leqq F_m\left(\frac{k}{\sqrt{1 + n\theta^2}}\right). \tag{5.8}$$

It follows that

$$\beta(3) = \int_0^\infty g_{mn}(\bar{r})\beta_0(\sqrt{n}\,A_2\bar{r})\,d\bar{r} \leqq \int_0^\infty g_{mn}(\bar{r})F_m\left(\frac{\sqrt{n}\,A_2\bar{r}}{\sqrt{1 + n\theta^2}}\right) d\bar{r}. \tag{5.9}$$

(5.9) is then expanded in a Taylor's series with remainder by the method of Section 3.

To obtain a lower bound, we have from elementary probability considerations $P(u < k \text{ or } v < k) = P(u < k) + P(v < k) - P(u < k, v < k)$, or

$$P(u < k, v < k) = P(u < k) + P(v < k) - P(u < k \text{ or } v < k). \tag{5.10}$$

Since the last term in (5.10) cannot exceed unity we have

$$P(u < k, v < k) = \beta_0(k) \geqq P(u < k) + P(v < k) - 1.$$

But

$$P(u < k) = P(v < k) = F_m\left(\frac{k}{\sqrt{1 + n\theta^2}}\right).$$

TABLE VII

*Bounds on* $\beta_0(3)$

| $\theta$ | $m = 5$ | | $m = 10$ | |
|---|---|---|---|---|
| | $n = 5$ | $n = 10$ | $n = 5$ | $n = 10$ |
| 0 | 1.00 | 1.00 | .98 .99 | .98 .99 |
| .5 | .87 .94 | | | |
| 1.5 | | .00 .12 | .00 .06 | .00 .01 |
| 2.0 | .00 .14 | .00 .05 | .00 .01 | |
| 2.5 | .00 .07 | .00 .02 | | |
| 3.0 | .00 .04 | .00 .01 | | |

| $\theta$ | $m = 15$ | | $m = 20$ | |
|---|---|---|---|---|
| | $n = 5$ | $n = 10$ | $n = 5$ | $n = 10$ |
| 0 | .97 .98 | .97 .98 | .96 .98 | .96 .98 |
| .75 | | | .00 .26 | .00 .04 |
| 1.0 | .00 .13 | .00 .02 | .00 .06 | |
| 1.25 | | | .00 .01 | |
| 1.5 | .00 .01 | | | |

$m = 25$

| $\theta$ | $n = 5$ | $n = 10$ |
|---|---|---|
| 0 | .95 .97 | .95 .97 |
| .25 | .80 .91 | |
| .75 | .00 .18 | .00 .02 |
| 1.00 | .00 .03 | |

Therefore

$$\beta_0(k) \geqq 2F_m\left(\frac{k}{\sqrt{1 + n\theta^2}}\right) - 1. \tag{5.11}$$

The corresponding lower bound on $\beta(3)$ becomes

$$\beta(3) \geqq \int_0^\infty g_{mn}(\bar{r}) \left[2F_m\left(\frac{\sqrt{n}\, A_2 \bar{r}}{\sqrt{1 + n\theta^2}}\right) - 1\right] d\bar{r}, \tag{5.12}$$

which is again evaluated by a Taylor expansion. The results given in Tables VII and VIII do not include certain intermediate values of $\theta$ in the range 0.25 to 1.50 because the lower bounds (5.11) and (5.12) are very weak in this range.

TABLE VIII

*Bounds on $\beta(3)$*

| $\theta$ | $m = 5$ | | $m = 10$ | |
|---|---|---|---|---|
| | $n = 5$ | $n = 10$ | $n = 5$ | $n = 10$ |
| 0 | .99 | 1.00 | .97 .99 | .98 .99 |
| .5 | .81 .99 | | | |
| 1.5 | | .00 .12 | .00 .06 | .00 .01 |
| 2.0 | .00 .14 | .00 .05 | .00 .01 | |
| 2.5 | .00 .07 | .00 .02 | | |
| 3.0 | .00 .04 | .00 .01 | | |

| $\theta$ | $m = 15$ | | $m = 20$ | |
|---|---|---|---|---|
| | $n = 5$ | $n = 10$ | $n = 5$ | $n = 10$ |
| 0 | .96 .98 | .96 .98 | .94 .97 | .95 .98 |
| .75 | | | .00 .27 | .00 .04 |
| 1.0 | .00 .13 | .00 .02 | .00 .06 | |
| 1.25 | | | .00 .01 | |
| 1.5 | .00 .01 | | | |

$m = 25$

| $\theta$ | $n = 5$ | $n = 10$ |
|---|---|---|
| 0 | .94 .97 | .94 .97 |
| .25 | .80 .90 | |
| .75 | .00 .18 | .00 .02 |
| 1.00 | .00 .03 | |

An important feature of Tables VII and VIII is the close agreement between corresponding values of $\beta_0(3)$ and $\beta(3)$. Our ignorance of $\sigma$ is of little consequence so far as these results are concerned.

**6. Bounds on the remainder term.** In each of the Taylor expansions used in Section 5, there is a remainder term of the form

$$R = \int_0^\infty \frac{b_{p+1}(\xi)}{(p+1)!} (\bar{r} - d_2)^{p+1} g_{mn}(\bar{r})\, d\bar{r}.$$

In order to determine a bound on $R$, we note that

$$|R| \leqq \int_0^\infty \frac{|b_{p+1}(\xi)|}{(p+1)!} |\bar{r} - d_2|^{p+1} g_{mn}(\bar{r})\, d\bar{r}$$

$$\leqq \frac{\text{Max} |b_{p+1}(\xi)|}{(p+1)!} \int_0^\infty |\bar{r} - d_2|^{p+1} g_{mn}(\bar{r})\, d\bar{r}.$$

If we restrict $p$ to be an odd integer, we have

$$|R| \leqq \frac{\text{Max} |b_{p+1}(\xi)|}{(p+1)!} \mu_{p+1}(\bar{r}). \tag{6.1}$$

In practically all cases given in the preceding tables, a sufficient number of terms of the series (3.7) is used to insure that the bound (6.1) does not exceed 0.005. In isolated instances where this is not possible, terms up to and including $b_6/6!\, \mu_6(\bar{r})$ are employed.

**7. Acknowledgment.** The author is indebted to Professor E. G. Olds for suggesting this problem and to both Professor Olds and Mr. R. B. Murphy for their kind guidance and encouragement in the course of study.

REFERENCES

[1] *ASTM Manual on Quality Control of Materials*, American Society for Testing Materials, Philadelphia, Pa., 1951.

[2] H. O. Hartley and E. S. Pearson, "Moment constants for the distribution of range in normal samples," *Biometrika*, Vol. 38 (1951), p. 463.

[3] F. E. Grubbs, "Sample criteria for testing outlying observations," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 27–58.

[4] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, 1946.

# THE USE OF PREVIOUS EXPERIENCE IN REACHING STATISTICAL DECISIONS[1]

By J. L. Hodges, Jr., and E. L. Lehmann

*University of California, Berkeley, and University of Chicago and Stanford University*

**1. Summary.** Instead of minimizing the maximum risk it is proposed to restrict attention to decision procedures whose maximum risk does not exceed the minimax risk by more than a given amount. Subject to this restriction one may wish to minimize the average risk with respect to some guessed a priori distribution suggested by previous experience. It is shown how Wald's minimax theory can be modified to yield analogous results concerning such restricted Bayes solutions. A number of examples are discussed, and some extensions of the above criterion are briefly considered.

**2. Introduction.** Among various possible approaches to the problem of defining a best decision procedure we may mention the following two extremes.

(i) *The Bayes solution.* If the unknown parameter $\theta$ is a random variable distributed according to a known probability distribution $\lambda$, and if $R_\delta(\theta)$ denotes the risk function of the decision procedure $\delta$, we simply minimize with respect to $\delta$ the average risk $\int R_\delta(\theta)\, d\lambda(\theta)$.

(ii) *The minimax principle.* Here one focuses attention on the maximum of the risk function and wishes to minimize $\sup_\theta R_\delta(\theta)$. The reader may consult Wald [1] for definitions and examples of these terms.

Of these two methods of treating the problem the first one assumes complete knowledge of the a priori distribution, an assumption that is usually not satisfied in practice. Even if extensive past experience is available, it will in most cases be difficult to exclude the possibility of some change in conditions. On the other hand, the minimax principle forces us to act as if $\theta$ were following a particular a priori distribution, the one least favorable to us, even though we may feel pretty sure that actually $\theta$ is distributed in quite a different manner. Thus it would seem that the minimax principle is suitable, if at all, only in situations characterized by a complete absence of past experience or other sources of knowledge concerning $\theta$.

The situations occurring in practice usually lie between the two extremes just described. On the one hand, one does frequently have a good idea as to the range of $\theta$, and as to which values in this range are more or less likely. On the other hand, such information cannot be expected to be either sufficiently precise or sufficiently reliable to justify complete trust in the Bayes approach.

The purpose of the present paper is to discuss an approach to the problem of

[1] Work sponsored in part by the Office of Naval Research.

optimal decisions that utilizes the available information but at the same time provides a safeguard in case this information is not correct.

Suppose that the maximum risk of the minimax procedure is $\underline{C}$. This is the smallest possible value for the maximum of a risk function. But we may be willing to tolerate a somewhat bigger maximum risk $C_0 > \underline{C}$ if, in case the guess at $\lambda$ has been a good one, there results a substantial decrease in the average risk. This leads to the

DEFINITION. The procedure $\delta_0$ is said to be a *restricted Bayes solution* with respect to the a priori distribution $\lambda$ and subject to the restriction

$$R_\delta(\theta) \leqq C_0 \text{ for all } \theta, \tag{2.1}$$

if it minimizes $\int R_\delta(\theta)\, d\lambda(\theta)$ among all procedures satisfying (2.1).

This definition takes into account two aspects of the risk function, its supremum and its average with result to the distribution $\lambda$. Within a certain range of values each of these can be improved at the expense of the other. The proper balance between the two, that is, the value we select for $C_0$, depends on the confidence we have in $\theta$'s actually following a distribution close to $\lambda$ and on the decrease in $\int R_\delta\, d\lambda$ that can be achieved by further increasing $C_0$.

To the above approach can also be given the following slightly different form. Instead of setting an upper bound for the risk, we specify a constant $0 \leqq \rho_0 \leqq 1$ and minimize

$$\rho_0 \int R_\delta(\theta)\, d\lambda(\theta) + (1 - \rho_0) \sup_\theta R_\delta(\theta). \tag{2.2}$$

Here it is $\rho_0$ that indicates the confidence we have in $\lambda$. The two principles are clearly equivalent; for if $\delta_0$ minimizes (2.2) and if $\sup R_{\delta_0}(\theta) = C_0$, then $\delta_0$ is a restricted Bayes solution, and the converse also holds.

The formulations given here may be applicable also to games played against an opponent rather than against Nature. This would be the case if one believed from past experience that the opponent is likely to make certain mistakes. One could then take advantage of these and still protect oneself in case the opponent has improved.

**3. Restricted Bayes solutions.** The principal aim of the present section is to obtain sufficient conditions for a decision procedure to be a restricted Bayes solution. For this purpose the modified problem mentioned at the end of the last section turns out to be the more natural one to consider; the results concerning restricted Bayes solutions follow as immediate corollaries. All of the theorems of this section will be simple generalizations of the corresponding results in Wald's minimax theory.

THEOREM 1. *Let $\nu_0$ be a distribution for which there exists a constant* $0 < \rho_0 \leqq 1$

*and a distribution $\mu_0$ such that*

$$\nu_0 = \rho_0\lambda + (1 - \rho_0)\mu_0 . \tag{3.1}$$

*Then if the Bayes solution $\delta_0$ of $\nu_0$ satisfies*

$$\int R_{\delta_0}(\theta)\, d\mu_0(\theta) = \sup_\theta R_{\delta_0}(\theta), \tag{3.2}$$

*the procedure $\delta_0$ minimizes*

$$\rho_0 \int R_\delta(\theta)\, d\lambda(\theta) + (1 - \rho_0) \sup R_\delta(\theta). \tag{3.3}$$

PROOF. Let $\delta$ be any procedure. Then

$$\rho_0 \int R_\delta\, d\lambda + (1 - \rho_0) \sup R_\delta(\theta) \geqq \rho_0 \int R_\delta\, d\lambda + (1 - \rho_0) \int R_\delta\, d\mu_0$$

$$\geqq \rho_0 \int R_{\delta_0}\, d\lambda + (1 - \rho_0) \int R_{\delta_0}\, d\mu_0 = \rho_0 \int R_{\delta_0}\, d\lambda + (1 - \rho_0) \sup R_{\delta_0}(\theta).$$

COROLLARY. *Under the assumptions of Theorem 1 suppose that* $\sup R_{\delta_0}(\theta) = C_0$. *Then $\delta_0$ minimizes $\int R_\delta\, d\lambda$ among all procedures satisfying $R_\delta(\theta) \leqq C_0$ for all $\theta$.*

Since a distribution $\nu_0$ with the required properties does not always exist, we state the following generalization of Theorem 1. *If $\nu_i$ is a sequence of a priori distributions with $\nu_i = \rho_0\lambda + (1 - \rho_0)\mu_i$, and if $\delta_i$ are the associated Bayes solutions, then*

$$\lim_{i\to\infty} \int R_{\delta_i}\, d\mu_i = \sup R_{\delta_0}(\theta) \tag{3.4}$$

*is a sufficient condition for $\delta_0$ to minimize* (3.3).

Before proceeding with this development let us discuss briefly the decomposition (3.1). For a given pair of distributions $\lambda$, $\nu$ consider the totality of numbers $0 \leqq \rho \leqq 1$ such that $\nu = \rho\lambda + (1 - \rho)\mu$ for some $\mu$. It is easily seen that this set is a closed interval $0 \leqq \rho \leqq \rho_\lambda$, and we shall call $\rho_\lambda$ the $\lambda$-*component* of $\nu$. It is of interest to note that under the conditions of Theorem 1 we have $\rho_0 = \rho_\lambda$ unless $C_0$ is the minimax risk. For if $\rho_0$ is not the $\lambda$-component of $\nu_0$ it follows that $\mu_0$ must have a positive $\lambda$-component. Thus any point of increase of $\lambda$ is also one of $\mu_0$ and we have $\int R_{\delta_0}\, d\lambda = C_0$. But by Theorem 1 this is possible only if $C_0$ is the minimax risk.

As in the minimax theory the distribution $\nu_0$ of Theorem 1 is "least favorable" for the statistician, in a sense made precise in

THEOREM 2. *Under the assumptions of Theorem 1, the distribution $\nu_0$ maximizes the Bayes risk among all a priori distributions $\nu$ that permit a representation*

$$\nu = \rho\lambda + (1 - \rho)\mu \text{ with } \rho \geqq \rho_0 .$$

PROOF. Let $\nu = \rho\lambda + (1 - \rho)\mu$ be any distribution with $\rho \geqq \rho_0$. Then

$$\int R_{\delta_0}\, d\nu_0 = \rho_0 \int R_{\delta_0}\, d\lambda + (1 - \rho_0) \sup R_{\delta_0}(\theta) \geqq \rho \int R_{\delta_0}\, d\lambda$$

$$+ (1 - \rho) \sup R_{\delta_0}(\theta) \geqq \rho \int R_{\delta_0}\, d\lambda + (1 - \rho) \int R_{\delta_0}\, d\mu$$

$$\geqq \int R_{\delta_\nu}(\theta)\, d\nu(\theta).$$

Thus $\delta_0$ is a maximum solution for Nature (in Wald's interpretation of a decision problem as a two-person zero sum game whose players are Nature and the statistician), if she is restricted to a priori distributions whose $\lambda$-component is at least $\rho_0$.

From the proof of Theorem 2 it is seen that $\nu_0$ not only maximizes the Bayes risk but also—and this is a slightly stronger result—the restricted Bayes risk, that is, the quantity $\inf \int R_\delta\, d\lambda$ when $\delta$ is restricted by the condition $R_\delta(\theta) \leqq C_0$ for all $\theta$.

In Theorem 1 we gave sufficient conditions for a procedure $\delta_0$ to be a restricted Bayes solution, and the distribution $\nu_0$ in terms of which these conditions were formulated was further characterized in Theorem 2. We must still prove the existence of a distribution with the desired properties, at least for some class of decision problems. This is easily done, along the lines of Wald's proof of Theorem 3.10 of [1], under the following assumptions.

ASSUMPTION 1.

$$\inf_\delta \sup_\mu \left[\rho_0 \int R_\delta\, d\lambda + (1 - \rho_0) \int R_\delta\, d\mu\right]$$

$$= \sup_\mu \inf_\delta \left[\rho_0 \int R_\delta\, d\lambda + (1 - \rho_0) \int R_\delta\, d\mu\right].$$

This states that the decision problem is strictly determined when Nature is restricted to distributions with $\lambda$-components $\geqq \rho_0$.

ASSUMPTION 2. *There exists a least favorable distribution $\mu_0$, that is a distribution that maximizes*

$$\inf_\delta \left[\rho_0 \int R_\delta\, d\lambda + (1 - \rho_0) \int R_\delta\, d\mu\right].$$

ASSUMPTION 3. *There exists a decision procedure $\delta_0$ that minimizes*

$$\sup_\mu \left[\rho_0 \int R_\delta\, d\lambda + (1 - \rho_0) \int R_\delta\, d\mu\right].$$

It follows from Wald's work that 1 and 3 hold, for example, when Conditions 3.1–3.6 of [1] are satisfied, while in general 2 requires the stronger Condition

(3.7). All these conditions are satisfied, for example, when we are dealing with discrete or absolutely continuous distributions, the problem is nonsequential, the loss function is bounded and the parameter space is compact.

To prove our result, we note that

$$\sup_{\mu}\left[\rho_0 \int R_\delta \, d\lambda + (1-\rho_0)\int R_\delta \, d\mu\right] = \rho_0 \int R_\delta \, d\lambda + (1-\rho_0)\sup_{\theta} R_\delta(\theta).$$

Therefore, we have by assumption

$$\rho_0 \int R_{\delta_0} \, d\lambda + (1-\rho_0)\sup R_{\delta_0}(\theta) = \inf_{\delta}\left[\rho_0 \int R_\delta \, d\lambda + (1-\rho_0)\int R_\delta \, d\mu_0\right],$$

and hence

$$\rho_0 \int R_{\delta_0} \, d\lambda + (1-\rho_0)\sup R_{\delta_0}(\theta) = \rho_0 \int R_{\delta_0} \, d\lambda + (1-\rho_0)\int R_{\delta_0} \, d\mu_0.$$

But this implies, when $\rho_0 < 1$, that

$$\sup R_{\delta_0}(\theta) = \int R_{\delta_0} \, d\mu_0.$$

On the other hand, the result is true vacuously when $\rho_0 = 1$.

**4. A continuity theorem.** We shall next consider the dependence of the restricted Bayes solution on the restricting quantity $C_0$. The main result will be a continuity theorem that permits us, at least in some cases, to characterize the topological structure of the restricted Bayes solutions.

Throughout this section we shall make the following two assumptions whose validity was proved by Wald (Theorems 3.1 and 3.2a of [1]) under very general conditions.

ASSUMPTION A. *The space of decision functions is compact, that is, every sequence $\delta_i$ of decision procedures possesses a subsequence $\delta_{i_j}$ that converges to some decision function $\delta^*$. Here convergence means what Wald calls "regular convergence"* (see pp. 65–66 of [1]).

ASSUMPTION B. *If $\delta_i \to \delta^*$, then for every distribution $\nu$*

$$\int R_{\delta_i} \, d\nu \to \int R_{\delta^*} \, d\nu.$$

This convergence of integrals implies the pointwise convergence of the risk functions, as may be seen by letting the distributions $\nu$ degenerate at single points $\theta$. In some cases (for example, if the loss function is bounded) the two notions of convergence are equivalent.

Let $\bar{C} = \sup R_{\delta_\lambda}(\theta)$ be the maximum risk of the unrestricted Bayes solution (for simplicity assume that $\delta_\lambda$ is unique; otherwise we would put $\bar{C} = \inf_{\delta_\lambda} \sup_\theta R_{\delta_\lambda}(\theta)$, the inf taken with respect to all Bayes solutions), and let $\underline{C}$ be the maximum risk of the minimax procedure. Then $\underline{C} \leqq \bar{C}$ and we may exclude the

case $\underline{C} = \bar{C}$ as trivial. For any $\underline{C} \leqq C \leqq \bar{C}$ let $\delta_\lambda(C)$ be an associated restricted Bayes solution, and $r_\lambda(C) = \int R_{\delta_\lambda(C)}\, d\lambda$ the corresponding restricted Bayes risk. Concerning $r_\lambda(C)$ we have the

LEMMA. *The function $r_\lambda(C)$ is convex, continuous, and strictly decreasing.*

PROOF.

(i) $r_\lambda(C)$ is obviously nonincreasing. If $C' < C''$, $0 < \gamma < 1$ and

$$\delta = \gamma\delta_\lambda(C') + (1 - \gamma)\delta_\lambda(C''),$$

then

$$\int R_\delta\, d\lambda = \gamma r_\lambda(C') + (1 - \gamma)\, r_\lambda(C'')$$

and

$$\sup R_\delta(\theta) \leqq \gamma C' + (1 - \gamma)C'',$$

from which the convexity of $r_\lambda(C)$ follows.

(ii) From well known properties of convex functions, the continuity of $r_\lambda(C)$ is now obvious except at the point $C = \underline{C}$. But let $C_i \downarrow \underline{C}$, let $\delta_i$ be the corresponding restricted Bayes solutions and $\delta_{i_j}$ the convergent subsequence guaranteed by Assumption A. If $\delta_{i_j} \to \delta^*$, we have $r_\lambda(C_i) = \int R_{\delta i_j}\, d\lambda \to \int R_{\delta *}\, d\lambda$. By monotonicity $r_\lambda(C) \geqq \int R_{\delta *}\, d\lambda$. Also, since $R_{\delta i_j}(\theta) \to R_{\delta *}(\theta)$ for each $\theta$, we have that $\sup R_{\delta *}(\theta) \leqq \lim C_{i_j} = \underline{C}$, so that $\sup R_{\delta *}(\theta) = \underline{C}$. But this implies $r_\lambda(C) \leqq \int R_{\delta *}\, d\lambda$ and hence $r_\lambda(C) = \int R_{\delta *}\, d\lambda = \lim r_\lambda(C_{i_j})$.

(iii) Before proving that $r_\lambda(C)$ is strictly decreasing we shall now show that for any $\underline{C} < C < \bar{C}$ we have

$$\sup_\theta R_{\delta_\lambda(C)}(\theta) = C.$$

Suppose that $\sup R_{\delta_\lambda}(\theta) < C$. If $\bar{C} < \infty$ there exists $0 < \gamma < 1$ so that $\delta = \gamma\delta_\gamma + (1 - \gamma)\delta_\lambda(C)$ still satisfies $\sup R_\delta(\theta) \leqq C$, but we would have $\int R_\delta\, d_\lambda < \int R_{\delta_\lambda(C)}\, d\lambda$. If $\bar{C} = \infty$, then $C_i \to \bar{C}$ implies that $\int R_{\delta_\lambda(C_i)}\, d\lambda \to \int R_{\delta_\lambda(\bar{C})}\, d\lambda$ and the same argument applies.

(iv) Strict monotonicity of $r_\lambda(C)$ is an obvious consequence of (iii). For let $C' < C''$ and suppose that $r_\lambda(C') = r_\lambda(C'')$. Then $\delta_\lambda(C')$ would be a solution not only corresponding to $C'$ but also to $C''$ in contradiction to (iii).

We can now state a closure and continuity theorem.

THEOREM 3. *If $\{\delta_i\}$ is a sequence of restricted Bayes solutions converging regularly to $\delta^*$, then $\delta^*$ is a restricted Bayes solution, and $\sup_\theta R_{\delta *}(\theta) = \lim_{i\to\infty} \sup_\theta R_{\delta_i}(\theta)$.*

PROOF. Let $r_i$, $r^*$ be the Bayes risks of $\delta_i$, $\delta^*$ respectively. By Assumption

B, $r_i \to r^*$. By the lemma, we can conclude that the sequence $C_i = \sup_\theta R_{\delta_i}(\theta)$ is convergent; denote $\lim C_i$ by $C_0$. Clearly $C^* = \sup R_{\delta^*}(\theta) \geqq C_0$, and since for each $\theta$ $R_{\delta_i^*}(\theta) = \lim R_{\delta_i}(\theta) \leqq C_0$, we conclude $C^* \leqq C_0 = C^*$. The lemma now assures $\delta^*$ to be a restricted Bayes solution.

COROLLARY. *The set of risk functions corresponding to restricted Bayes solutions is closed with respect to the convergence of Assumption B.*

PROOF. Let $R_i$ be the risk functions corresponding to restricted Bayes solutions $\delta_i$, and let $R_i \to R^*$. By the compactness Assumption A, we may extract a subsequence $\delta_{i_j}$ which converges regularly to some $\delta^*$. By Theorem 3, $\delta^*$ is a restricted Bayes solution, whose risk function $R^*$ is the limit of $R_{\delta_{i_j}}$, and hence of $R_{\delta_i}$.

**5. Extensions.** We shall now mention briefly some extensions of the notion of a restricted Bayes solution. To give a first simple example, it may happen that we have ideas concerning the range of the parameter but not concerning a possible distribution over this range. Thus it may be known that $\theta \varepsilon \Omega$ and may further be indicated by other considerations that actually $\theta \varepsilon \omega$ where $\omega \subset \Omega$. We may then, subject to the condition $R_\delta(\theta) \leqq C_0$ for all $\theta \varepsilon \Omega$, wish to minimize $\sup_{\theta \varepsilon \omega} R_\delta(\theta)$. For example, when testing the hypothesis that three means $\theta_1$, $\theta_2$, $\theta_3$ are equal we may believe that the most likely alternatives are such that $\theta_1 < \theta_2 < \theta_3$ without however definitely being able to exclude the other possibilities.

We can get further refinements as follows. Let $\Lambda_0 \subset \Lambda_1 \subset \cdots \subset \Lambda_{r-1} \subset \Lambda_r$ be a nested set of families of distributions. We may be certain that the true distribution of $\theta$ is an element of $\Lambda_r$ ($\Lambda_r$ may of course contain all distributions that assign probability 1 to a single point), fairly sure that it lies in $\Lambda_{r-1}$, still believe that is in $\Lambda_{r-2}$, etc. Accordingly we could select a decreasing sequence of constants $C_r > C_{r-1} > \cdots > C_1$, each $C_i$ greater than the restricted minimax risk at this stage, and minimize $\sup_{\lambda \varepsilon \Lambda_i} \int R_\delta \, d\lambda$ subject to the condition $\sup_{\lambda \varepsilon \Lambda_i} \int R_\delta \, d\lambda \leqq C_i$ for $i = 1, \cdots, r$.

There is an extension of Theorem 1 to the present case that we give as an example of how the theory of the earlier sections generalizes.

THEOREM 4. *Suppose there exists $\rho_0 > 0$ and a distribution*

$$\nu_0 = \rho_0 \lambda_0 + \sum_{i=1}^{r} \rho_i \lambda_i, \qquad \rho_i \geqq 0, \qquad \sum_{i=0}^{r} \rho_i = 1,$$

*with $\lambda_i \varepsilon \Lambda_i$ and such that the Bayes solution $\delta_0$ of $\nu_0$ satisfies*

$$\sup_{\lambda \varepsilon \Lambda_i} \int R_{\delta_0} \, d\lambda = C_i = \int R_{\delta_0} \, d\lambda_i \qquad \text{for } i = 1, \cdots, r \tag{1}$$

*and*

$$\sup_{\lambda \varepsilon \Delta_0} \int R_{\delta_0} \, d\lambda = \int R_{\delta_0} \, d\lambda_0. \tag{2}$$

*Then* $\delta_0$ *minimizes* $\sup_{\lambda \in \Lambda_0} \int R_\delta \, d\lambda$ *subject to*

$$\sup_{\lambda \in \Lambda_i} \int R_\delta \, d\lambda \leqq C_i \qquad \text{for } i = 1, \cdots, r. \tag{3}$$

PROOF. Let $\delta$ be any procedure satisfying (3). Then

$$\sum_{i=0}^{r} \rho_i \sup_{\lambda \in \Lambda_i} \int R_\delta \, d\lambda \geqq \sum \rho_i \int R_\delta \, d\lambda_i \geqq \sum \rho_i \int R_{\delta_0} \, d\lambda_i = \sum \rho_i \sup_{\lambda \in \Lambda_i} \int R_\delta \, d\lambda.$$

But for $i = 1, \cdots, r$ we have

$$\sup_{\lambda \in \Lambda_i} \int R_\delta \, d\lambda \leqq C_i = \sup_{\lambda \in \Lambda_i} \int R_{\delta_0} \, d\lambda$$

and the conclusion follows.

Of the conditions, (2) of course becomes vacuous in case $\Lambda_0$ contains only a single distribution.

**6. Examples.** We conclude the paper by discussing a number of examples which serve to illustrate the ideas and theorems. As might be expected, it is more difficult to obtain explicit results with the restricted Bayes approach than with the logically simpler minimax or Bayes principles. In fact, the minimax principle has been successfully applied, in most cases, by guessing the answer and then verifying it through use of the specialized form of Theorem 1. It appears that the restricted Bayes solutions are often much less simple in their mathematical structure than the minimax solutions, and will accordingly be much harder to guess. We suspect that the widespread application of the restricted Bayes approach would require numerical methods in combination with theoretical results of the kind given above.

EXAMPLE 1. Let $X = (X_1, X_2, \cdots, X_n)$ be random variables having the joint density function $p_\theta(x) = c(\theta)b(x)e^{t(x)q(\theta)}$, where $\theta$ is an unknown real parameter and $q(\theta)$ is a monotonely increasing function of $\theta$. This form includes many of the distributions frequently used in statistics, such as the Poisson, binomial, and normal with known mean or variance. Consider the problem of testing, at a given level of significance, the hypothesis $H: \theta = \theta_0$ against the alternatives $\theta \neq \theta_0$. We are indifferent to alternatives $\theta_1 \leqq \theta \leqq \theta_2$, where $\theta_1 < \theta_0 < \theta_2$. As risk function, we shall use the probability of false acceptance. Thus, the risk is $1 - \beta(\theta)$, where $\beta(\theta)$ is the power function. Suppose finally that past experience or other considerations suggest that, should the hypothesis $\theta = \theta_0$ be false, then one of the alternatives $\theta \geqq \theta_2$ is true.

It is known [2] that an essentially complete class of tests consists in those of the form $w(k)$: reject $H$ if $t(x) \leqq k$ or $t(x) \geqq f(k)$, where $f(k)$ is determined by the given level of significance. We may restrict attention to these tests. Denote the power of $w(k)$ at alternative $\theta$ by $\beta_k(\theta)$. If we had complete confidence in the presumption that either $\theta = \theta_0$ or $\theta \geqq \theta_2$, we should seek to maximize the power against the latter alternatives, with no regard to the power against the alternatives $\theta \leqq \theta_1$.

This leads to the known uniformly most powerful one-sided test $w(-\infty)$. This test is obtainable as the Bayes test corresponding to any a priori distribution for $\theta$ which assigns all of the probability to values $\theta \geqq \theta_2$. However, the power of $w(-\infty)$ tends to 0 as $\theta$ tends to its lower limit, so we may get very poor performance if our presumption is not correct. At the other extreme, if we placed no reliance on the presumption, we might seek the minimax test. It is easy to see that the power functions all have unique minima, and are continuous in $k$ for fixed $\theta$, whence the minimax test will be $w(\underline{k})$, where $\underline{k}$ is determined by the condition $\beta_{\underline{k}}(\theta_1) = \beta_{\underline{k}}(\theta_2)$. This test then gives no better protection against large alternatives than against small, and thus makes no use at all of our presumption that large alternatives are the ones to fear.

The restricted Bayes approach suggests a compromise, under which we would seek to maximize $\inf_{\theta \geqq \theta_2} \beta(\theta)$ subject to $\inf_\theta \beta(\theta) \geqq C$. It can be shown that, for $-\infty < k \leqq \underline{k}$, $\inf_{\theta \geqq \theta_2} \beta_k(\theta) = \beta_k(\theta_2)$, while $\inf_\theta \beta_k(\theta)$ is attained at a finite value of $\theta \leqq \theta_1$. From these facts the Lemma of [2] enables us to conclude, for $k' < k < k''$, $k \leqq \underline{k}$, that $\inf_{\theta \geqq \theta_2} \beta_k(\theta) > \inf_{\theta \geqq \theta_2} \beta_{k''}(\theta)$, while $\inf_\theta \beta_k(\theta) > \inf_\theta \beta_{k'}(\theta)$. It follows that for every $-\infty < k \leqq \underline{k}$, $w(k)$ is a restricted Bayes solution in the sense given. The same result holds if we assign to $\theta$, any a priori distribution under which $P(\theta \geqq \theta_2) = 1$, and apply the restricted Bayes principle in the narrower sense of Section 2.

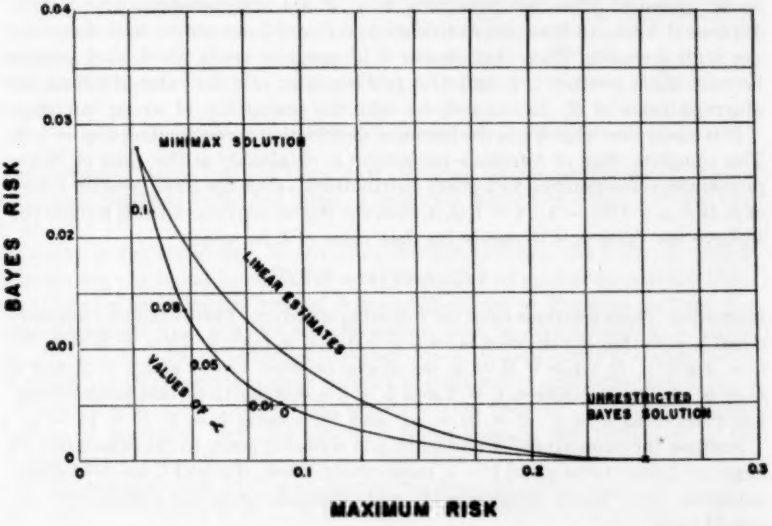
It is easily seen that $C(k) = \inf_\theta \beta_k(\theta)$ is a continuous, monotonely increasing function of $k$ in the interval $-\infty \leqq k \leqq \underline{k}$. The tests $w(k)$, $k \leqq \underline{k}$, thus provide essentially unique restricted Bayes solutions for all possible values of $C$. In this example, our principle provides nothing essentially new, since the admissible tests already form only a one-parameter family.

EXAMPLE 2. Consider next the binomial random variable $X$ = number of successes on $n$ independent trials, each having the probability $p$ of success, and the problem of estimating $p$ from $X$. Take as loss function the square of the error of estimate. Suppose that a theoretical examination of the experimental situation reveals considerations of symmetry which suggest that $p = \frac{1}{2}$. We wish to design an estimate which will take advantage of this theory, in the sense of providing small risk at $p = \frac{1}{2}$, but will not place complete reliance in the theory, in the sense of giving some control over the risk for all values of $p$. This problem falls within the framework of the restricted Bayes approach, if we give to $p$ the a priori distribution $\lambda$ which assigns all of its probability to the value $\frac{1}{2}$. Our objective is then to minimize the risk at $\frac{1}{2}$, subject to a given maximum $C$ for the risk over $0 \leqq p \leqq 1$.

Denote the estimate corresponding to a value $x$ of $X$ by $\delta(x)$, and the risk involved in using this estimate by $R_\delta(p) = E[\delta(X) - p]^2$. The ordinary Bayes estimate is $\bar{\delta}(x) = \frac{1}{2}$ for every $x$, since this estimate reduces the risk at $\frac{1}{2}$ to 0, while every other estimate has positive risk at $\frac{1}{2}$. The maximum risk of $\bar{\delta}$ is $\bar{C} = \frac{1}{4}$, attained at $p = 0$ and $p = 1$. The minimax estimate for this problem is known (see [3]) to be $\underline{\delta}(x) = \dfrac{\sqrt{n}}{\sqrt{n}+1}\dfrac{X}{n} + \dfrac{1}{2(\sqrt{n}+1)}$. This estimate has the constant risk $\underline{C} = \frac{1}{4}(\sqrt{n}+1)^2$.

In getting the restricted Bayes estimates, we first observe that we need only consider those estimates which possess the symmetry property $\delta(x) + \delta(n - x) = 1$. For, if $\delta(x)$ is any estimate, the estimate $\delta'(x) = 1 - \delta(n - x)$ will have the same maximum risk and the same Bayes risk as $\delta$, for any a priori distribution which is symmetric about $p = 1/2$. Thus, the estimate $\frac{1}{2}(\delta(x) + \delta'(x)$, which possesses the symmetry property cited, will have the same Bayes risk as $\delta$ and no larger maximum risk.

Consider the distribution $\mu_1$ which assigns probability $1/2$ to each of the points 0 and 1. The Bayes solution $\delta_1$ corresponding to the a priori distribution $\rho\lambda + (1 - \rho)\mu_1$ is easily found to be of the form: $\delta_1(0) = 1 - \delta_1(n)$, $\delta_1(x) = 1/2$ for $0 < x < n$. By virtue of Theorem 2, $\delta_1$ is a restricted Bayes estimate if it achieves




its maximum risk only at 0 and 1. But this is true for $\bar{R}(p) = (p - \frac{1}{2})^2$, and hence by continuity (Theorem 3) must hold true for some interval $(C_1, \bar{C})$ of values of $C$. In fact, $C_1$ may be calculated from the condition $R'_1(0) = 0$. Again using the continuity of the variation of the risk function with $C$, we see that for some interval $(C_2, C_1)$, the distribution $\mu_2$ will consist of probability $1/2$ at points $\alpha$ and $1 - \alpha$, the value of $\alpha$ varying continuously with $C$. At $C = C_2$, the risk function will develop further maxima, requiring a modification of the form of the distribution $\mu$, and so forth. But in any case, the distribution $\mu$ must have only finitely many points of positive probability, since the risk function must achieve its maximum at all such points, is a polynomial in $p$, and is not constant except for $C = \underline{C}$.

The computing program outlined in the preceding paragraph was carried out numerically for the case $n = 4$, leading to the relation between $C$ and $R(1/2)$

shown in the graph. The restricted Bayes estimates are (except for the extremes of minimax and Bayes estimates) not linear functions of $X$. For comparison, the graph also shows the minimum risk at $p = 1/2$ attainable for given maximum risk using only linear estimates. We see that, over a wide range of values of $C$, the use of nonlinear estimates permits nearly a 50 per cent reduction in the risk at $p = 1/2$.

EXAMPLE 3. As an illustration of the application of our principle to a finite decision problem, consider the following genetic situation. We are concerned with the inheritance of a simple Mendelian trait. Two individuals $F$ and $G$ are crossed to produce $n$ progeny, each of whom is then crossed with a known hybrid to produce a single offspring. Of the $n$ third-generation individuals, $R$ are found to be recessive, while the remaining $n - R$ are either dominant or hybrid. Suppose it is known from other evidence that $F$ and $G$ are neither both dominant nor both recessive. Then the number $k$ of recessive genes which they possess between them is either 1, 2, or 3. Our problem is to infer the value of $k$ from the observed value of $R$. As our risk, we take the probability of wrong inference.

It is easily seen that $R$ has the binomial distribution corresponding to $p = k/8$. The complete class of inference procedures is obtainable as the class of Bayes procedures corresponding to a priori distributions $\lambda$ over the three possible values of $p$. If $P(p = i/8) = \lambda_i$, $i = 1, 2, 3$, then the Bayes solution, when $R$ is observed to have the value $r$, is to decide for that value of $k$ for which

$$\lambda_k \binom{n}{r} (k/8)^r (1 - k/8)^{n-r}$$

is greatest. These solutions have the following structure. There exist two numbers $a$ and $b$, such that our decision is for $k = 1$ if $R < a$, for $k = 2$ if $a < R < b$, for $k = 3$ if $b < R$, while if $R = a$, we choose between $k = 1$ and $k = 2$, and if $R = b$, we choose between $k = 2$ and $k = 3$ according to certain probabilities. Let $P(\text{choosing } k = 1 \mid R = a) = \pi_1$, and $P(\text{choosing } k = 2 \mid R = b) = \pi_2$.

Suppose the proportions of dominant and recessive genes in the population at large are known to be $\mu$ and $1 - \mu$, respectively. Then, if $F$ and $G$ are bred under panmixia (i.e., chosen independently and randomly from the population), we should have

$$P(k = 1) : P(k = 2) : P(k = 3) = 2(1 - \mu)^2 : 3\mu(1 - \mu) : 2\mu^2.$$

The Bayes solution corresponding to this distribution for $k$ would be the reasonable decision procedure to use if we were sure that $F$ and $G$ had been bred under panmixia. If we placed no reliance at all in the panmixia hypothesis, we might prefer to employ the minimax solution, which is characterized by equal probabilities of error corresponding to the three possible values of $k$. If we placed some reliance in the panmixia hypothesis, but not complete reliance, we might impose a limit on the permissible probability of error for any value of $k$, and subject to this limit seek to minimize the average probability of error under panmixia.

The minimax, Bayes, and restricted Bayes solutions are easy to obtain numerically. For illustration, suppose $n = 20$ and $\mu = 0.8$. The three solutions are as follows.

| *Solution* | $a$ | $b$ | $\pi_1$ | $\pi_2$ |
|---|---|---|---|---|
| Minimax | 3 | 7 | 0.482 | 0.142 |
| Bayes | 0.5 | 4.5 | ... | ... |
| Restricted Bayes, maximum probability of error = 1/2 | 2 | 5 | 0.868 | 0.828 |

The performance characteristic properties of the three solutions may be conveniently compared in tabular form.

| *Solution* | *Maximum probability of error* | *Average probability of error under panmixia* |
|---|---|---|
| Minimax | 0.354 | 0.354 |
| Restricted Bayes | 0.500 | 0.239 |
| Bayes | 0.931 | 0.234 |

The restricted Bayes solution loses little efficiency as compared with the Bayes solution if panmixia holds, but gives considerably better protection if it does not hold.

It may be remarked that, in any finite decision problem, the restricted Bayes solution may be found by means of a finite number of applications of the Neyman-Pearson fundamental lemma. Suppose there are $m$ decisions, so that the risks associated with any decision procedure may be represented as a vector $(r_1, r_2, \cdots, r_m)$ in $m$-dimensional Euclidean space. Except in the trivial situation in which the Bayes solution itself satisfies the restrictive condition that $r_i \leqq C$ for every $i = 1, 2, \cdots, m$, we may conclude from part (iii) of the proof of the Lemma of Section 3, that the restricted Bayes solution will have $r_i = C$ for at least one $i$. Let $I(C)$ be the set of $i$ such that $r_i = C$. The restricted Bayes solution corresponding to $C$ may then be obtained by minimizing $\sum \lambda_j r_j$, subject to $r_i = C$ for all $i$ in $I(C)$. If we compute these minimizing solutions for all sets $I$, using the Neyman-Pearson lemma, we may select the restricted Bayes solution from among them by inspection.

## REFERENCES

[1] A. Wald, *Statistical Decision Functions*, John Wiley and Sons, 1950.

[2] E. L. Lehmann, "On families of admissible tests," *Annals of Math. Stat.*, Vol. 18 (1947), pp. 97–104.

[3] J. L. Hodges, Jr. and E. L. Lehman, "Some problems in minimax point estimation," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 182–197.

# ON THE EXISTENCE OF LEAST FAVORABLE DISTRIBUTIONS[1]

By E. L. Lehmann

*Stanford University and University of California, Berkeley*

**1. Summary.** Sufficient conditions for the existence of a least favorable distribution were given by Wald in his work on general decision theory. It is shown here that for problems of hypothesis testing and more generally for multiple decision problems involving a finite number of decisions, the result holds under a much weaker restriction than Wald's assumption of a compact parameter space.

**2. Introduction.** In his general theory of decision functions Wald has given conditions under which there exists a least favorable distribution for Nature. One of these, which is rather restrictive and not required for many parts of the theory, is condition (3.7) of [1], a compactness assumption on the parameter space. This condition can, of course, in general not be dispensed with; there are many examples of simple statistical decision problems for which a least favorable distribution does not exist. However, as we shall show below, for certain special classes of problems the compactness assumption is not required.

The following is a typical example. Suppose that $\theta$ is a location parameter and that we are interested in testing $H: \theta \leqq \theta_0$ against $K: \theta \geqq \theta_1$. If a least favorable distribution does not exist this means essentially that positive probability is assigned to at least one of the points $\theta = \pm\infty$. But it is intuitively clear that by doing this, Nature would be playing into the hands of the statistician, since for sufficiently large $|\theta|$ it is very easy to determine whether or not $H$ is true. In fact one would expect Nature to do better if the above strategy $S$ were replaced by the (conditional) distribution for $\theta$ obtained from $S$, given that no probability is placed on $\theta = \pm\infty$.

In the next section we shall make this argument precise and use it to establish the existence of a least favorable distribution for certain hypothesis testing and multiple decision problems. In particular our condition is satisfied for many problems of testing a composite hypothesis against a simple alternative. This shows that a method given in [2] for determining the most powerful test against a simple alternative is applicable in most problems of the kind usually considered.

**3. Least favorable distributions for problems of hypothesis testing.** In the present section we shall consider the hypothesis $H$ that the probability density of a random variable $X$ is $f(x)$ with $f \in \mathfrak{F}$, where $\mathfrak{F}$ is some given class of densities, and the class of alternatives $K$ that the density of $X$ is $g_\theta(x)$ with $\theta \in \Omega$. We shall assume that the sample space $\mathfrak{X}$ is Euclidean and that $\theta = (\theta_1, \cdots, \theta_s)$, where $\Omega$ is a Borel set in $R_s$. Since throughout we shall be concerned with dis-

[1] Work sponsored in part by the office of Naval Research.

tributions over $\Omega$ we may, by defining these distributions to be zero outside of $\Omega$, assume without loss of generality that $\Omega$ is the full Euclidean space $R_s$. We assume further that for almost all $x$, $g_\theta(x)$ is continuous in $\theta$ and measurable in $\Omega \times \mathfrak{X}$.

For the present we consider the usual asymmetric formulation of the problem in which a level of significance $\alpha$ is assigned and it is desired to maximize the minimum power of the test over $K$. In this situation, as is pointed out on page 133 of [1], assumptions (3.1) – (3.6) of [1] hold, and we have, as a simple consequence of Theorem 3.12 of [1],

LEMMA 3.1. *There exists a level-$\alpha$ test $\phi$ of $H$ that maximizes the minimum power against $K$, and a sequence of distributions $\lambda_i$ over $\Omega$ such that*

$$\lim_{i\to\infty} \int E_\theta \phi(X)\, d\lambda_i(\theta) = \inf_{\theta\in\Omega} E_\theta \phi(X) \tag{3.1}$$

*and*

$$\lim_{i\to\infty} \beta_i = \inf_{\theta\in\Omega} E_\theta \phi(X), \tag{3.2}$$

*where $\beta_i$ is the power of the most powerful level-$\alpha$ test for testing $H$ against the simple alternative*

$$\int g_\theta(x)\, d\lambda_i(\theta).$$

Now by Helly's compactness theorem there exists a subsequence of the sequence $\{\lambda_i\}$, which without loss of generality we shall take to be the original sequence, such that

$$\lambda_i \to \gamma\mu \tag{3.3}$$

for all continuity intervals of $\mu$, where $\mu$ is again a distribution and $0 \leqq \gamma \leqq 1$. If here we could assume that $\gamma = 1$ it would follow that $\mu$ is least favorable (see the proof of Theorem 3.14 of [1]). We shall now show that $\mu$ is least favorable by proving that either $\gamma = 1$ or $\lim \beta_i = 1$ under the following

ASSUMPTION A. *Given any $\epsilon > 0$ and any closed bounded set $\omega \subset \Omega$ there exists a Borel set $S$ in the sample space such that*

$$\begin{aligned} P(\bar{S} \mid f) &\leqq \epsilon && \text{for all } f \in \mathfrak{F}, \\ P(\bar{S} \mid g_\theta) &\leqq \epsilon && \text{for all } \theta \in \omega, \end{aligned} \tag{3.4}$$

*and*

$$P(S \mid g_\theta) \to 0 \text{ uniformly as } \theta_1^2 + \cdots + \theta_s^2 \to \infty. \tag{3.5}$$

Clearly Assumption A is satisfied if both $\mathfrak{F}$ and $\Omega$ are compact, but as we shall show later the assumption also holds in many cases in which (3.7) of [1] is not satisfied.

The proof of our result is based on the following two lemmas.

LEMMA 3.2. *Let $\omega$ be any closed bounded set in $\Omega$ and let $\epsilon > 0$. Then there exists a Borel set $S$ in $\mathfrak{X}$ and an integer $i_0$ such that*

$$\int_{\bar{\omega}} P(\bar{S} \mid \theta)\, d\lambda_i(\theta) \geqq 1 - \gamma - \epsilon \qquad \text{for } i \geqq i_0, \tag{3.6}$$

*and even*

$$\int_{\bar{\omega}} E_\theta[\phi(X) c_S(X)]\, d\lambda_i(\theta) \geqq 1 - \gamma - 2\epsilon \qquad \text{for } i \geqq i_0, \tag{3.7}$$

*where $\phi$ is the test referred to in Lemma 3.1 and where $c_S(x)$ is the characteristic set function of $\bar{S}$.*

Here and in what follows the notation $P(\bar{S} \mid \theta)$ and $E_\theta$ indicates that the probability or expectation is computed with respect to the probability density $g_\theta(x)$.

LEMMA 3.3. *For any $\epsilon > 0$ there exists a continuity interval $\omega$ of $\mu$ such that for any test $\phi$*

$$\int_\omega E_\theta \phi(X)\, d\lambda_i(\theta) \geqq \gamma \int_\Omega E_\theta \phi(X)\, d\mu(\theta) - \epsilon \tag{3.8}$$

*for sufficiently large $i$.*

Before proving these lemmas we shall show that they imply $\gamma = 1$ or $\lim \beta_i = 1$ and hence the existence of a least favorable distribution. Let $\phi$ be the test referred to in Lemma 3.1 and let

$$\lim_{i \to \infty} \int E_\theta \phi(X)\, d\lambda_i(\theta) = \beta. \tag{3.9}$$

Given any $\epsilon > 0$ let $\omega$ be a bounded subset of $\Omega$ for which (3.8) holds. Then

$$\int_\Omega E_\theta \phi(X)\, d\lambda_i(\theta) = \int_\omega E_\theta \phi(X)\, d\lambda_i(\theta) + \int_{\bar{\omega}} E_\theta \phi(X)\, d\lambda_i(\theta),$$

where for sufficiently large $i$ the first term is $\geqq \gamma \int_\Omega E_\theta \phi(X)\, d\mu(\theta) - \epsilon$ by (3.8) and the second term is $\geqq 1 - \gamma - 2\epsilon$ by (3.7), so that

$$\beta \geqq (1 - \gamma) + \gamma \int_\Omega E_\theta \phi(X)\, d\mu(\theta) \geqq \int_\Omega E_\theta \phi(X)\, d\mu(\theta). \tag{3.10}$$

Here the last inequality is strict and therefore contradicts (3.1), unless either

$$\gamma = 1 \text{ or } \int_\Omega E_\theta \phi(X)\, d\mu(\theta) = 1.$$

But in the latter case $\beta = 1$ by (3.10), so that

$$\int_\Omega E_\theta \phi(X)\, d\mu(\theta) = \inf E_\theta \phi(X)$$

and $\mu$ would be least favorable.

PROOF OF LEMMA 3.2. Let $S$ be a set guaranteed by Assumption A for which

$$P(\bar{S} \mid \theta) \leqq \epsilon/2 \qquad \text{for all } \theta \in \omega, \tag{3.11}$$

$$P(\bar{S} \mid f) \leqq \epsilon \qquad \text{for all } f \in \mathfrak{F}, \tag{3.12}$$

and such that

$$P(S \mid \theta) \to 0 \text{ uniformly as } \theta_1^2 + \cdots + \theta_s^2 \to \infty. \tag{3.13}$$

From (3.3) and (3.13) it follows that

$$\int_\Omega P(\bar{S} \mid \theta)\, d\lambda_i(\theta) \geqq 1 - \gamma - \epsilon/2 \qquad \text{for } i \geqq i_0, \tag{3.14}$$

and (3.6) is an immediate consequence of (3.14) and (3.11).

Suppose now that (3.7) did not hold. Then there would exist a subsequence of $\{\lambda_i\}$, which we may again take to be the full sequence such that

$$\lim_{i\to\infty} \int_{\bar{\omega}} E_\theta[\phi(X)c_\beta(X)]\, d\lambda_i(\theta) = \delta < 1 - \gamma. \tag{3.15}$$

We define a new test $\phi^*$ by

$$\phi^*(x) = \begin{cases} 1 & \text{if } x \in \bar{S}, \\ \dfrac{\alpha - \epsilon}{\alpha}\,\phi(x) & \text{if } x \in S. \end{cases} \tag{3.16}$$

Then for any $f \in \mathfrak{F}$ we have, because of (3.12)

$$E(\phi^*(X) \mid f) \leqq \frac{d - \epsilon}{\alpha}\,\alpha + P(\bar{S} \mid f) \leqq \alpha, \tag{3.17}$$

so that $\phi^*$ is again a level-$\alpha$ test. Also, for all $x$,

$$\phi^*(x) \geqq \frac{\alpha - \epsilon}{\alpha}\,\phi(x) + c_\beta(x)[1 - \phi(x)], \tag{3.18}$$

and hence

$$\begin{aligned} \int E_\theta\phi^*(X)\, d\lambda_i(\theta) \geqq \frac{\alpha - \epsilon}{\alpha} \int E_\theta\phi(X)\, d\lambda_i(\theta) \\ + \int E_\theta\{c_\beta(X)[1 - \phi(X)]\}\, d\lambda_i(\theta). \end{aligned} \tag{3.19}$$

The first term on the right-hand side of (3.19) tends to $\beta(\alpha - \epsilon)/\alpha$ as $i \to \infty$. On the other hand the second term is greater than or equal to

$$\int_{\bar{\omega}} P_\theta(\bar{S})\, d\lambda_i(\theta) - \int_{\bar{\omega}} E_\theta[c_\beta(X)\phi(X)]\, d\lambda_i(\theta),$$

which for sufficiently large $i$ exceeds $(1 - \gamma - \epsilon) - (\delta + \epsilon)$ by (3.15) and (3.6). This shows that for sufficiently large $i$

$$\int E_\theta \phi^*(X)\, d\lambda_i(\theta) > \beta + \frac{1-\gamma-\delta}{2}$$

and thus contradicts (3.2), since it implies that for sufficiently large $i$ $\phi^*$ is more powerful than the most powerful test for testing $H$ against $\int g_\theta(x)\, d\lambda_i(\theta)$.

PROOF OF LEMMA 3.3. Let $\omega$ be a continuity interval of $\mu$, and so large that

$$\mu(\bar{\omega}) \leqq \epsilon/2. \tag{3.20}$$

Then by the Helly-Bray theorem, as $i \to \infty$

$$\int_\omega g_\theta(x)\, d\lambda_i(\theta) \to \gamma \int_\omega g_\theta(x)\, d\mu(\theta),$$

and hence by Fatou's lemma

$$\liminf_i \int \left[\phi(x) \int_\omega g_\theta(x)\, d\lambda_i(\theta)\right] dx \geqq \gamma \int \left[\phi(x) \int_\omega g_\theta(x)\, d\mu(\theta)\right] dx.$$

Interchanging the order of integration we then have for $i$ sufficiently large

$$\int_\omega E_\theta \phi(X)\, d\lambda_i(\theta) \geqq \gamma \int_\omega E_\theta \phi(X)\, d\mu(\theta) - \epsilon/2 \geqq \gamma \int_\Omega E_\theta \phi(X)\, d\mu(\theta) - \epsilon$$

by (3.20).

We can sum up the work of this section in

THEOREM 3.1. *If $g_\theta(x)$ is continuous in $\theta$ and if Assumption A holds, then for the problem of testing $H$ against $K$ there exists a least favorable distribution for $\theta$.*

As an illustration, let $X_1, \cdots, X_n$ be a sample from a distribution with density $f(x-\theta)$ and consider testing $H: \theta \leqq \theta_0$ against $\theta \geqq \theta_1$. Here we could clearly take for $S$ a set of the form $X_1 \leqq c$, so that Assumption A is satisfied.

**4. Testing against a simple alternative.** As an application of Theorem 3.1 we now consider the problem of testing a composite hypothesis against a simple alternative. Let the hypothesis state that the density of $X$ is $f_\theta(x)$, $\theta \in \Omega$, and let the alternative be $g(x)$, where we assume as before that the sample space is Euclidean and that $\theta = (\theta_1, \cdots, \theta_s)$ with $\Omega$ being the full Euclidean space $R_s$.

ASSUMPTION B. *For every closed bounded set $S$ in the sample space, $P(S \mid f_\theta) \to 0$ uniformly as $\theta_1^2 + \cdots + \theta_s^2 \to \infty$.*

THEOREM 4.1. *If $f_\theta(x)$ is a continuous function of $\theta$ for almost all $x$, then under Assumption B there exists a least favorable distribution over $\Omega$.*

PROOF. Theorem 3.1 clearly applies if we interchange $H$, $\alpha$ with $K$, $\beta$. We therefore only need to show that Assumption B implies Assumption A with $f$ and $g$ interchanged. That is, we need to prove that Assumption B implies the existence of a set $S$ such that

$$P(\bar{S} \mid g) \leqq \epsilon, \tag{4.1}$$

$$P(\bar{S} \mid f_\theta) \leqq \epsilon \qquad \text{for all } \theta \in \omega, \tag{4.2}$$

and

$$P(S \mid f_\theta) \to 0 \text{ as } \theta_1^2 + \cdots + \theta_s^2 \to \infty. \tag{4.3}$$

It was pointed out by Scheffé [3] that when $g_\theta(x)$ is continuous in $\theta$, the convergence $\theta_i \to \theta$ implies

$$\int_S g_{\theta_i}(x)\,dx \to \int_S g_\theta(x)\,dx \text{ uniformly for all Borel sets } S. \tag{4.4}$$

Therefore if $\omega$ is bounded and closed, the associated set of distributions is compact with respect to the convergence definition (4.4). Let $\epsilon > 0$. Then we can find a finite number of points in $\omega$, say $\theta_1, \cdots, \theta_r$, such that every $\theta \in \omega$ is within $\frac{1}{2}\epsilon$ of one of these $\theta_i$ in the sense that $|P(S \mid \theta) - P(S \mid \theta_i)| \leqq \frac{1}{2}\epsilon$ for all Borel sets $S$. It follows that there exists a closed bounded set $S$ for which (4.1) and (4.2) hold and (4.3) then follows from Assumption B.

ASSUMPTION C. *There exist $s$ statistics $Y_i = h_i(X)$, $i = 1, \cdots, s$, such that the $h_i$ are continuous, the marginal distribution of $Y_i$ depends only on $\theta_i$, and $Y_i$ tends to $+\infty$ or $-\infty$ as $\theta_i \to +\infty$ or $-\infty$.*

LEMMA 4.1. *A sufficient condition for Assumption B to hold is Assumption C.*

PROOF. Let $S$ be a closed bounded set, and let $a_i = \min_S h_i(x)$, $b_i = \max_S h_i(x)$. Then $P_\theta(X \in S) \leqq P_{\theta_i}(a_i \leqq Y_i \leqq b_i)$ for $i = 1, \cdots, s$. Now if $\theta_1^2 + \cdots + \theta_s^2 > C$ we must have $\theta_i^2 > C/s$ for at least one value of $i$. Given $\epsilon > 0$, let $C$ be so large that

$$P_{\theta_i}(a_i \leqq Y_i \leqq b_i) < \epsilon \qquad \text{if} \qquad \theta_i^2 > \frac{C}{s} \qquad \text{for all } i.$$

Then $P_\theta(X \in S) \leqq \epsilon$ provided $\theta_1^2 + \cdots + \theta_s^2 > C$.

As an example let $X_1, \cdots, X_n$ be a sample from a distribution with unknown location and scale parameter $\xi$ and $\eta$. Let $\theta_1 = \xi$, $\theta_2 = \log \eta$, and take $Y_1 = X_1/|X_2 - X_1|$, $Y_2 = \log|X_2 - X_1|$. Then Assumption C is seen to be satisfied, so that for testing a hypothesis of this type against a simple alternative, a least favorable distribution exists.

Theorem 4.1 answers a problem raised by [2], where in Theorem 1 a method is given for proving that a test is most powerful for testing a composite hypothesis against a simple alternative. Combining Theorem 4.1 with Theorem 3.10 of [1], we see that this method is always applicable when Assumption B is satisfied.

While it is quite likely that Assumption B can be weakened somewhat, it is perhaps of interest to point out that some condition of this type is required, which restricts the behavior of the distributions as $\theta_1^2 + \cdots + \theta_s^2 \to \infty$. As an example consider the following situation in which $\Omega$ is denumerable and the densities are given by

$$H: f_i(x) = \frac{1}{2\sqrt{2\pi}}\left[e^{-\frac{1}{2}(x-i)^2} + e^{-\frac{1}{2}(x-(1/i))^2}\right], \qquad i = m, m+1, \cdots,$$

$$K: g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

We shall show that if $m$ is sufficiently large the most powerful level-$\alpha$ test rejects when $x \leqq b$, where $b$ is determined by

$$\lim_{i\to\infty} \int_{-\infty}^{b} f_i(x)\,dx = \tfrac{1}{2} \int_{-\infty}^{b} g(x)\,dx = \alpha. \tag{4.5}$$

First one checks easily that for large $i$ $\alpha(i) = \int_{-\infty}^{b} f_i(x)\,dx$ is an increasing function of $i$. Let $m$ be such that $\alpha(i)$ is increasing for $i \geqq m$. Then the proposed test is of size $\alpha$ and power $2\alpha$ by (4.5). But if $\beta$ is the power of the most powerful test $\phi$ we clearly have

$$\beta \leqq 2 \limsup_{i\to\infty} \int f_i(x)\,\phi(x)\,dx \leqq 2\alpha,$$

which proves that the region of rejection $x \leqq b$ is most powerful. If a least favorable distribution $\lambda$ over $\Omega$ existed, the test $x \leqq b$ would be a Bayes solution for $\lambda$ (Theorem 3.9 of [1]), that is, would be most powerful for testing $\int f_i(x)\,d\lambda(i)$ against $g(x)$. But an application of the Neyman-Pearson fundamental lemma shows immediately that no Bayes solution is of this form.

It may be worth pointing out that the conditions of the present paper are really of a quite different nature from those given by Wald in his general theory. Thus they do not even imply that the parameter space is weakly compact in the sense of (3.1) of [1]. As an example, consider a random variable $X$ the distribution of which depends on an unknown location parameter $\theta$. Let $H$ denote the hypothesis $\theta \leqq \theta_0$ and suppose that the simple alternative is $\theta_1 > \theta_0$. Let $\{\lambda_i\}$ be any sequence of distributions over $H$, and let $\phi(x)$ be any test. Then weak compactness of $H$ would imply the existence of a subsequence $\{\lambda_{i_j}\}$ and a distribution $\lambda_0$ over $H$ such that

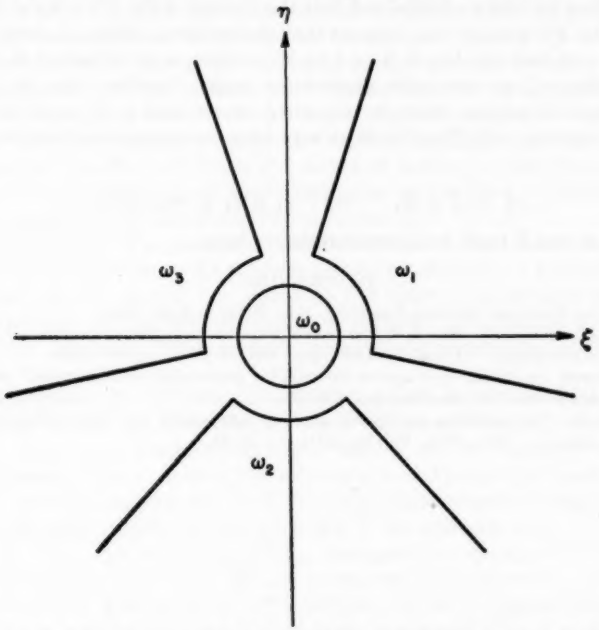
$$\limsup_{j\to\infty} \int E_\theta\,\phi(x)\,d\lambda_{i_j}(\theta) \leqq \int E_\theta\,\phi(x)\,d\lambda_0(\theta). \tag{4.6}$$

But let $\phi(x)$ be the characteristic set function of the set $x \leqq C$ and let $\lambda_i(\theta)$ assign probability 1 to the point $\theta = -i$. Then the left-hand side of (4.6) is 1, while for any distribution $\lambda_0$ the right-hand side is $<1$, so that $H$ is not weakly compact. On the other hand, Assumption B is satisfied.

**5. Some extensions.** We shall now indicate briefly how the results of Section 3 may be extended to somewhat more general decision problems. As a first generalization consider the problem of hypothesis testing, formulated as in Section 3 except that the loss function need no longer be simple. Suppose that acceptance of the hypothesis when the true distribution is given by $g_\theta(x)$, $\theta \in \Omega$ results in a loss $W(\theta)$, and that $V(f)$ is the loss resulting from rejection of $H$ when $f$ is the true density. We shall assume that $V$ and $W$ are bounded and that we are interested in minimizing $\sup_\theta W(\theta)E_\theta[1 - \phi(x)]$ among all tests $\phi$ for which $V(f)E[\phi(x) \mid f] \leqq \alpha$ for all $f \in \mathfrak{F}$.

Let Assumption A′ be obtained from Assumption A by replacing $P(\bar{S} \mid g_\theta)$ and $P(S \mid g_\theta)$ in (3.4) and (3.5) by $W(\theta)P(\bar{S} \mid g_\theta)$ and $W(\theta)P(S \mid g_\theta)$ respectively, and $P(\bar{S} \mid f)$ in (3.4) by $V(f)P(\bar{S} \mid f)$. Then Assumption A′ implies the existence of a least favorable distribution over $\Omega$. Only the obvious changes are required in the proof of Theorem 3.1.

As an immediate further generalization one next obtains an analogous result for the following two-decision problem. The parameter space is partitioned into two set $\omega_1 \cup \omega_2$. The loss is zero if $\theta \in \omega_i$ and decision $d_i$ is taken; it is $W_i(\theta)$ if




decision $d_i$ is taken incorrectly. Let us denote the risk function of a decision procedure $\delta$ as usual by $R_\delta(\theta)$ and suppose that if $\delta$ is the minimax solution

$$\sup_{\omega_i} R_\delta(\theta) = \gamma .$$

Then the overall minimax problem is equivalent (see [4]) to the problem $\Pi_1$ of minimizing $\sup_{\theta \in \omega_1} R_\delta(\theta)$ subject to $R_\delta(\theta) \leqq \gamma$ for all $\theta \in \omega_2$, and also to the problem $\Pi_2$ obtained from $\Pi_1$ by interchanging 1 and 2. If a least favorable distribution over $\omega_i$ exists for $\Pi_i$ $(i = 1, 2)$, then one exists for the overall problem. But each of the partial problems is of the type discussed at the beginning of this section.

The whole argument extends without much difficulty to multiple decision

problems involving a finite number of decisions and having a bounded loss function. We shall not give the details but simply state the result. Let $\omega_1, \cdots, \omega_K$ be the regions in the parameter space in which the decisions $d_1, \cdots, d_K$ are appropriate so that the loss $W(\theta, d_i) = 0$ if $\theta \varepsilon \omega_i$. Then we consider the two-decision problems $\Pi_i$ $(i = 1, \cdots, K)$, where the decision lies only between $\theta \varepsilon \omega_i$ and $\theta \varepsilon \Omega - \omega_i$. A least favorable distribution exists over $\Omega$ provided each of the problems $\Pi_i$ satisfies Assumption A'.

As an example consider the case that $X_1, \cdots, X_m$; $Y_1, \cdots, Y_n$ are $m + n$ independent random variables, and that the density of the $X$'s is $f(x - \xi)$ while that of the $Y$'s is $g(y - \eta)$. Suppose that the partition of $\Omega$ is as shown in the diagram and that the loss is 0 or 1 as the correct or an incorrect decision is taken. Then a least favorable distribution exists. To show that $\Pi_3$ satisfies Assumption A' suppose that the bounding rays $l_1$ and $l_2$ of $\omega_3$ are given by $\eta = m_1\xi$ and $\eta = m_2\xi$. Then for $\bar{S}$ we may take a set in the sample space of the form

$$x_1^2 + y_1^2 \geqq R, \qquad m_1 + c_1 \leqq y_1 \leqq m_2 x_1 + c_2,$$

where $R$, $c_1$ and $c_2$ must be chosen sufficiently large.

## REFERENCES

[1] A. Wald, *Statistical Decision Functions*, John Wiley & Sons, 1950.

[2] E. L. Lehmann and C. Stein, "Most powerful tests of composite hypotheses. I. Normal distributions," *Annals of Math. Stat.*, Vol. 19 (1948), pp. 495–516.

[3] H. Scheffé, "A useful convergence theorem for probability distributions," *Annals of Math. Stat.*, Vol. 18 (1947), pp. 434–438.

[4] C. Blyth, "On minimax statistical decision procedures and their admissibility," *Annals of Math. Stat.*, Vol. 22 (1951), pp. 22–42.

# SUFFICIENT STATISTICS AND SELECTION DEPENDING ON THE PARAMETER

By D. A. S. Fraser

*University of Toronto*

**1. Summary.** The concept of "functional sufficiency" or "f-sufficiency" for a class of density functions is introduced and conditions given under which it corresponds to sufficiency as defined by Halmos and Savage [2]. A minimal f-sufficient statistic is defined and shown to exist, and its construction is given.

The minimal f-sufficient statistic for a class of densities for which the region of positive density varies with the parameter is shown to be equivalent to the combination of a "statistic of selection" and the minimal f-sufficient statistic for a class of densities for which the region of positive density is fixed. The construction of sufficient statistics in this latter case subject to certain conditions of regularity has been treated by Koopman [1].

If the parameter is a parameter of selection from a fixed distribution, then the statistic of selection is the minimal f-sufficient statistic. If in addition the regions of positive density are monotone and indexed monotonely by a real parameter, then the statistic of selection is sufficient according to the Halmos and Savage definition. Three examples are given to illustrate the results.

**2. Introduction.** The following notation follows closely that introduced in [2] by Halmos and Savage. Let $X$ be a general space and $S$ a Borel class over it; then $(X, S)$ is called a measurable space. Since $X$ will be considered the space of a chance quantity, a statistic $T$ is a function over $X$, that is, a transformation from $X$ into a measurable space $(T, S')$. The transformation is measurable if the inverse transformation of the Borel sets $S'$ are elements of $S$.

Let $\mathfrak{M} = \{\mu\}$ be a class of probability measures over the space $X$. If $\mu$ and $\nu$ are two measures on $X$, $\nu$ is absolutely continuous with respect to $\mu$, in symbols $\nu \ll \mu$, if $\nu(E) = 0$ for every $E \in S$ for which $\mu(E) = 0$. A set of measures $\mathfrak{M}$ on $X$ will be called dominated if there exists a measure $\lambda$ on $X$ such that for each $\mu \in \mathfrak{M}$ we have $\mu \ll \lambda$, in notation $\mathfrak{M} \ll \lambda$.

In this paper we consider only dominated sets of measures $\mathfrak{M} \ll \lambda$. The Radon-Nikodym Theorem gives us the following statement:

*For any dominated set of measures $\mathfrak{M} \ll \lambda$, there exists a set of nonnegative real-valued functions over $X$, $\{f_\mu(x)\}$, such that*

$$\mu(E) = \int_E f_\mu(x)\, d\lambda(x)$$

*for every $\mu \in \mathfrak{M}$ and $E \in S$. The functions $f_\mu(x)$ are unique except on a set of $\lambda$-measure zero.* We shall refer to $f_\mu(x)$ as a density function relative to the fixed measure $\lambda$.

For any such class of measures and a corresponding set of densities, this paper presents theorems concerning sufficient statistics.

**3. Definition of sufficient statistic.** The concept of sufficient statistic was given a rigorous formulation in [2]. In words the definition is as follows.

*$T$ is sufficient for $\mathfrak{M}$ if there exists for almost all $(\mu T^{-1})$ values of $T$ a conditional probability function which is independent of $\mu \in \mathfrak{M}$.*

Corollary 1 to Theorem 1 in [2] gives an equivalent condition for a sufficient statistic:

*The statistic $T$ is sufficient for a dominated set of measures if and only if $f_\mu(x)$ factors in the form $f_\mu(x) = g_\mu(T)h(x)$ (except on a set of $\mu$-measure zero), where $g_\mu(T)$ is measurable, $g$ and $h$ are nonnegative, and $h$ and $gh$ are integrable with respect to $\lambda$.*

For the purposes of this paper we introduce the concept of functional sufficiency with respect to a class of density functions $\{f_\mu(x)\}$ over a fixed measure $\lambda$. For these density functions $\mu$ can be considered as a parameter or index of the family. *A function $T(x)$ is functionally sufficient ($f$-sufficient) for $\{f_\mu(x)\}$ if and only if $f_\mu(x) = g_\mu(T)h(x)$, where $g_\mu(T)$, $h(x)$ are nonnegative.* For the study of f-sufficiency the normalization and integrability conditions on $f_\mu(x)$ are not important; however, the property that $f_\mu(x) \geqq 0$ will be assumed throughout. *A function $T(x)$ is a minimal $f$-sufficient statistic*[1] *for $\{f_\mu(x)\}$ if no function of $T(x)$ (other than 1-1) is $f$-sufficient for $\{f_\mu(x)\}$.*

LEMMA 1. *For any $\{f_\mu(x)\}$ there exists a minimal $f$-sufficient statistic.*

PROOF. The minimal f-sufficient statistic will be given as a mapping of the elements of $X$ onto a class of disjoint subsets covering $X$. The mapping is from the elements of a set to the set itself. Thus we need only define the disjoint sets covering $X$. Define sets as follows:

$$E_0 = \{x \mid f_\mu(x) \text{ is independent of } \mu\},$$

$$E_{x'} = \{x \mid \exists h(x) \text{ such that } f_\mu(x) = h(x)f_\mu(x')\}.$$

By this procedure the whole space $X$ can be covered. Also a simple argument shows that the sets are disjoint.

We must now show that these sets provide us with an f-sufficient statistic. Let $T$ index the sets and $T(x)$ be the mapping from $x$ to the set containing $x$. Then we have either $f_\mu(x) = h_0(x)$, or $f_\mu(x) = h_T(x)f_\mu(x')$, where $x' = x'(T)$. But this implies $f_\mu(x) = g_\mu(T)h(x)$, and hence $T$ is f-sufficient.

For any function $T'(T)$ other than 1-1 there would exist at least two sets $E'$, $E''$ giving the same value to the function $T'$. Over $E'$, $E''$ the representation $f_\mu(x) = g_\mu(T')h(x)$ is impossible, since it would imply that the sets were identical. Thus $T$ is a minimal f-sufficient statistic.

[1] The concept of a minimal $f$-sufficient statistic was obtained independently of [4]. The partition of $X$ induced by the minimal $f$-sufficient statistic is essentially the partition produced by the operator $\vartheta$ introduced by Lehmann and Scheffé.

LEMMA 2. *There is a unique minimal f-sufficient statistic and it is a function of any other f-sufficient statistic.*

PROOF. Let $T(x)$ be any f-sufficient statistic: then $f_\mu(x) = g_\mu(T)h(x)$. Consider the sets $E$ corresponding to the partition of $X$ induced by the function $T(x)$. Within any one $E$ of the sets

$$f_\mu(x) = g_\mu(E)h(x).$$

Hence $E$ is contained in a set of the partition induced by the minimal f-sufficient statistic defined in Lemma 1 and therefore that minimal f-sufficient statistic is a function of $T(x)$. This with the definition of minimal f-sufficient statistic establishes uniqueness.

LEMMA 3. *Any sufficient statistic for a dominated set of measures $\mathfrak{M} \ll \lambda$ is an f-sufficient statistic for an equivalent set of densities (relative to $\lambda$).*

PROOF. Let $T$ be any sufficient statistic for $\mathfrak{M}$. By Corollary 1 to Theorem 1 in [2], there exists a set of densities $\{f_\mu(x)\}$ such that $f_\mu(x) = g_\mu(T)h(x)$ except on a set of $\mu$-measure zero. An equivalent set of densities is $\{g_\mu(T)h(x)\}$ and these factor for all values of $x$. Hence $T$ is f-sufficient for $\{f_\mu(x)\}$.

LEMMA 4. *An f-sufficient statistic $T$ for $\{f_\mu(x)\}$ is sufficient for $\{\mu\} = \mathfrak{M}$ if and only if $f_\mu(x) = g_\mu(T)h(x)$ where $T(x)$ is measurable, $g_\mu(T)$ is measurable, and $h$ and $gh$ are integrable with respect to $\lambda$.*

PROOF. Follows immediately from Corollary 1, Theorem 1 [2].

The structure of f-sufficiency, unlike sufficiency, is not invariant under changes in $\{f_\mu(x)\}$ which do not alter the set $\mathfrak{M}$. However for applications the conditions in Lemma 4 will usually be fulfilled, and consequently any procedure for obtaining f-sufficient statistics will assist the study of sufficient statistics.

**4. Selection.** Consider a set of densities $\{f_\mu(x)\}$ over a space $X$. Considering $\mu$ as an index or parameter of the class, we define a parameter of selection.

*$\mu$ is a "parameter of selection" for $\{f_\mu(x)\}$ if*

$$f_\mu(x) = \phi_{S_\mu}(x)f(x)g(\mu)\,,$$

*where $S_\mu = \{x \mid f_\mu(x) > 0\}$ and $\phi_S(x) = 0, 1$, according as $x \notin S, \in S$.*

If there is a class of measures $\mathfrak{M}$ which are essentially truncations of, or selections from a fixed measure, then there will exist an equivalent set of densities for which $\mu$ is a "parameter of selection."

We now define a "statistic of selection" for any set of densities $\{f_\mu(x)\}$. It is a function mapping $X$ into the space of subsets of $X$.

*$T(x)$, a statistic of selection, is defined by*

$$T(x) = \bigcap S_\mu\,,$$

*where the intersection is over all $\mu \in \mathfrak{M}^*(x) = \{\mu \mid x \in S_\mu\} = \{\mu \mid f_\mu(x) > 0\}$.*

We also define a characteristic function for sets $A$ and $B \subset X$:

$$\psi(A, B) = 0 \text{ if } A \not\subset B,$$
$$= 1 \text{ if } A \subset B.$$

**5. F-sufficiency and the statistic of selection.** A set of measures $\mathfrak{M}$ is said to be homogeneous if $\nu \ll \mu$ and $\mu \ll \nu$ for every pair $\mu, \nu \in \mathfrak{M}$. Thus a dominated set of measures with densities $\{f_\mu(x)\}$ is homogeneous if $S_\mu$ is independent of $\mu$. In such a case the Halmos and Savage criterion of sufficiency is simplified, and if working over the real space subject to certain continuity conditions on the $f_\mu(x)$ the problem of sufficiency has been treated by Koopman [1].

The theorem in this section reduces the problem of finding an f-sufficient (or minimal f-sufficient) statistic for a set of densities corresponding to a dominated set of measures to that of finding an f-sufficient (or minimal f-sufficient) statistic for the special type of densities (mentioned above) corresponding to a homogeneous set of measures. Application of these results for sufficiency then rests on Lemmas 3 and 4 and the fact that in applications sufficiency and f-sufficiency will usually be the same.

THEOREM 1. *For a class of densities $f_\mu(x)$ relative to a fixed measure $\lambda$ on $X$, the minimal f-sufficient statistic is the combination of the statistic of selection and the minimal f-sufficient statistic for a class of densities for which $S_\mu$ is independent of $\mu$ (the carrier independent of the parameter).*

PROOF. With no loss of generality we can assume $\bigcup S_\mu = X$, where the union is over all $\mu \in \mathfrak{M}$.

We first show the equivalence of the two characteristic functions $\phi_{S_\mu}(x)$ and $\psi(T(x), S_\mu)$, where $T(x)$ is the statistic of selection.

$$\phi_{S_\mu}(x) = 0, 1$$

implies

$$x \notin S_\mu, \qquad x \in S_\mu,$$

which implies

$$T(x) \not\subset S_\mu, \qquad T(x) \subset S_\mu,$$

which implies

$$\psi(T, S_\mu) = 0, 1.$$

Hence

$$\phi_{S_\mu}(x) \equiv \psi(T(x), S_\mu).$$

The density function $f_\mu(x)$ can be written as follows:

$$\begin{aligned} f_\mu(x) &= \phi_{S_\mu}(x) f_\mu(x) \\ &= \psi(T, S_\mu) f_\mu(x) \\ &= \psi(T, S_\mu) \bar{f}_\mu(x), \end{aligned} \tag{5.1}$$

where

$$\begin{aligned} \bar{f}_\mu(x) &= f_\mu(x) \text{ if } x \in S_\mu, \\ &= g_\mu(x) \text{ if } x \notin S_\mu. \end{aligned}$$

$g_\mu(x)$ is chosen everywhere greater than zero and of a functional form introduced later in the proof; it can as shown in the proof be chosen independent of $\mu$, that is $g_\mu(x) = g(x)$.

The functions $\bar{f}_\mu(x)$ are everywhere positive and hence the carrier $\bar{S}_\mu$ is independent of the parameter $\mu$. The choice of function $g_\mu(x)$ or $g(x)$ still has a high degree of arbitrariness, and will be made in application in such a way as to facilitate the determination of a minimal f-sufficient statistic for $\{\bar{f}_\mu(x)\}$.

Let $T'$, $T''$ be minimal f-sufficient statistics for $\{f_\mu(x)\}$, $\{\bar{f}_\mu(x)\}$, then

$$\begin{aligned} f_\mu(x) &= g_\mu(T')h(x), \\ \bar{f}_\mu(x) &= \bar{g}_\mu(T'')\bar{h}(x). \end{aligned} \tag{5.2}$$

Substituting (5.2) in (5.1), we obtain

$$f_\mu(x) = \psi(T, S_\mu)\bar{g}_\mu(T'')\bar{h}(x).$$

Thus $(T, T'')$ is an f-sufficient statistic for $\{f_\mu(x)\}$, and we need only establish that it is minimal.

For the last section of the proof, we shall need to know that $T(x)$ is constant valued for all $x$ for which $T'$ has a fixed value. This is true since $T$ is a function of $T'$:

$$T = \bigcap S_\mu$$

where the intersection is over all $\mu \in \mathfrak{M}^*(x) = \{\mu \mid f_\mu(x) > 0\} = \{\mu \mid g_\mu(T') > 0\}$.

As in Lemma 1, we study the functions $T'(x)$ and $(T(x), T''(x))$ by considering the partitions they induce on the space $X$. Any set of the partition is thus the set of points for which the function takes a constant value. For the partition induced by $(T, T'')$, we shall show that any set is also a set of the partition corresponding to $T'(x)$. Thus $T'$ and $(T, T'')$ can be put into 1-1 correspondence and hence $(T, T'')$ is minimal f-sufficient.

Let $E$ be any set of the partition of $X$ corresponding to $(T, T'')$; then

$$\begin{aligned} E_{x'} &= \{x \mid T(x) = T(x');\ \exists h(x) \ni \bar{f}_\mu(x) = h(x)\bar{f}_\mu(x') \text{ for all } \mu\} \\ &= \left\{x \,\middle|\, \begin{array}{l} T(x) = T(x');\ \exists h(x) \ni f_\mu(x) = h(x)f_\mu(x') \text{ for all } \mu \text{ for which} \\ S_\mu \supset T(x), \text{ and } g_\mu(x) = h(x)g_\mu(x') \text{ for all } \mu \text{ for which } S_\mu \not\supset T(x) \end{array}\right\} \\ &= \left\{x \,\middle|\, \begin{array}{l} T(x) = T(x');\ \exists h(x) \ni f_\mu(x) = h(x)f_\mu(x') \text{ for all } \mu, \text{ and} \\ g_\mu(x) = h(x)g_\mu(x') \text{ for all } \mu \text{ for which } S_\mu \not\supset T(x) \end{array}\right\} \qquad (5.3) \\ &= \left\{x \,\middle|\, \begin{array}{l} \exists h(x) \ni f_\mu(x) = h(x)f_\mu(x') \text{ for all } \mu, \text{ and } g_\mu(x) = h(x)g_\mu(x') \\ \text{for all } \mu \text{ for which } S_\mu \not\supset T(x) \end{array}\right\}. \qquad (5.4) \end{aligned}$$

(5.4) follows from (5.3) since it was shown above that $T(x) = T(x')$ is a consequence of $T'(x) = T'(x')$.

$$\begin{aligned} E_{x'} &= \left\{x \,\middle|\, \begin{array}{l} \exists h(x) \ni f_\mu(x) = h(x)f_\mu(x') \text{ for all } \mu, \text{ and } g_\mu(x) = \\ h(x)g_\mu(x') \text{ for all } \mu \text{ for which } S_\mu \not\supset T(x) \end{array}\right\} \\ &= \{x \mid \exists h(x) \ni f_\mu(x) = h(x)f_\mu(x') \text{ for all } \mu\}, \end{aligned}$$

if $g(x)$, which so far is only restricted to positive values, is given the following functional form:

$$g_\mu(x) = c_\mu(T) f_{\nu_\mu(T)}(x)$$

where $\nu_\mu(T(x))$ can be any element of $\mathfrak{M}$ for which $S_\nu \supset T(x)$. Thus $E$ is a set of the partition corresponding to $T'$. This completes the proof of Theorem 1.

As remarked in the proof, $g_\mu(x)$ can be chosen independent of $\mu$, as for example

$$g(x) = C(T) f_{\nu(T)}(x)$$

where $\nu(T(x))$ can be any element of $\mathfrak{M}$ for which $S_\nu \supset T(x)$.

In a recent paper [3] R. C. Davis has investigated sufficient statistics for density functions on $R^1$ satisfying certain regularity conditions and having the upper and lower extremities of the range of the distribution depend on the parameter. The proofs of necessity in Theorems 1 and 2 in [3] follow immediately from Lemma 4 and the theorem in this section by noting that in each case the statistic considered is equivalent to the statistic of selection.

**6. Parameter of selection.** In the special case in which $\mu$ acts as a parameter of selection for $\{f_\mu(x)\}$, the following theorem gives explicitly an f-sufficient statistic.

THEOREM 2. *If $\mu$ acts as a parameter of selection for $\{f_\mu(x)\}$ then the 'statistic of selection' is the minimal f-sufficient statistic for $\{f_\mu(x)\}$.*

PROOF. Since $\mu$ is a parameter of selection, $f_\mu(x)$ has the form

$$\begin{aligned} f_\mu(x) &= \phi_{S_\mu}(x) f(x) g(\mu) \\ &= \psi(T, S_\mu) f(x) g(\mu). \end{aligned}$$

Hence $T$ is f-sufficient. By Theorem 1, $T$ is minimal f-sufficient for $\{f_\mu(x)\}$.

**7. Real-valued parameter.** We now consider f-sufficiency when the parameter is real-valued and is a parameter of selection.

THEOREM 3. *If $\mu$ is a parameter of selection for $\{f_\mu(x)\}$, $\{S_\mu\}$ are Borel sets and ordered by $\subset$, and $S_\mu$ can be put in monotone 1-1 correspondence with a subset of $R^1$, then there exists a real-valued statistic which is sufficient and minimal f-sufficient for $\{f_\mu(x)\}$.*

PROOF. By Theorem 2 the statistic of selection is minimal f-sufficient, hence we want a real valued statistic $T^*$ which can be put into 1-1 correspondence with the statistic of selection. First we define a real valued parameter essentially equivalent to $\mu$.

By hypothesis we can choose $\theta(\mu) \in R^1$ such that $S_{\mu'} \subset S_{\mu''} \leftrightarrow \theta(\mu') \leqq \theta(\mu'')$ and $S_{\mu'} = S_{\mu''} \leftrightarrow \theta(\mu') = \theta(\mu'')$. Since $\mu$ is a parameter of selection, an equivalent parameter is $S_\mu$ and as is seen above $\theta$ is equivalent to $S_\mu$. Hence we use the parameter $\theta$ to index the densities.

We define $T^*(x) \in R^1$ as follows:

$$T^*(x) = \inf \theta \text{ for all } \theta \text{ having } x \in S_\theta .$$

Therefore

$$T^*(x) = \inf \theta \text{ for all } \theta \text{ having } T_x \subset S_\theta .$$

Thus $T^*$ is a function of $T$. If $T^*$ is also f-sufficient then it can be put into 1-1 correspondence with $T$ and is minimal f-sufficient. Define $\psi^*(T^*, \theta)$ by

$$\psi^*(T^*, \theta) = 1 \quad \text{if } T^* \leqq \theta,$$
$$= 0 \quad \text{if } T^* > \theta.$$

We now prove $\psi^*(T^*(x), \theta) \equiv \psi(T(x), S_\theta)$.

$$\psi^*(T^*(x), \theta) = 1$$

is equivalent to

$$T^*(x) \leqq \theta,$$

which is equivalent to

$$T(x) \subset S_\theta ,$$

that is,

$$\psi(T(x), S_\theta) = 1.$$

Hence

$$\psi^*(T^*(x), \theta) \equiv \psi(T(x), S_\theta).$$

The densities $f_\mu(x)$ have the following form:

$$\begin{aligned} f_\mu(x) &= \phi_{S_\theta}(x) \cdot f(x) \cdot g(\theta) \\ &= \psi(T(x), S_\theta) \cdot f(x) \cdot g(\theta) \\ &= \psi^*(T^*(x), \theta) \cdot f(x) \cdot g(\theta). \end{aligned}$$

Therefore $T^*(x)$ is f-sufficient and hence by previous argument is minimal f-sufficient. A more direct proof of this statement could have been made by establishing the 1-1 correspondence, but the above functional form for $f_\mu(x)$ is desirable for the remainder of the proof.

To prove that $T^*(x)$ is sufficient we need only verify that the conditions of Lemma 4 are fulfilled for the set of densities $\{f_\mu(x)\}$. Since the sets $S_\theta$ are Borel sets, $T^*(x)$ is a measurable function. Also $\psi^*(T^*, \theta)$ is a measurable function of $T^*$, and $f_\mu(x)$ is integrable. We now verify that $f(x)$ or a modification of $f(x)$ is integrable over $\bigcup S_\theta$. Take a monotone sequence $\{\theta_i\}$ such that $S_{\theta_i} \to \bigcup S_\theta$. Let

$$\int_{\theta_1} f(x)\, d\lambda(x) = e_1.$$

Choose $r(\theta_1) > 0$ such that

$$\int_{S_{\theta_2} - S_{\theta_1}} \frac{f(x)}{r(\theta_1)}\, d\lambda(x) = \tfrac{1}{2} e_1,$$

and similarly

$$\int_{S_{\theta_i}-S_{\theta_{i-1}}} \frac{f(x)}{r(\theta_{i-1})}\, d\lambda(x) = e_1/2^{i-1}.$$

Define

$$\begin{aligned} r(T^*) &= 1 && \text{if} \quad T^* \leqq \theta_1, \\ &= r(\theta_1) && \text{if} \quad \theta_1 < T^* \leqq \theta_2, \\ &= r(\theta_i) && \text{if} \quad \theta_i < T^* \leqq \theta_{i+1}. \end{aligned}$$

Then

$$\int_{S_{\theta_i}} \frac{f(x)}{r(T^*)}\, d\lambda(x) = 2e_1[1 - (\tfrac{1}{2})^i].$$

Therefore

$$\int_{US_\theta} \frac{f(x)}{r(T^*)}\, d\lambda(x) = 2e_1.$$

Therefore the modification of $f(x)$, that is, $f(x)/r(T^*)$, is integrable over $US_\theta$. The theorem follows by noting that the modification does not affect the factorization needed for sufficiency:

$$f_u(x) = \psi^*(T^*(x), \theta) \cdot g(\theta) \cdot r(T^*(x)) \cdot \frac{f(x)}{r(T^*(x))}.$$

**8. Examples.**

EXAMPLE 1. Consider the class of all uniform distributions on $R^1$ having range 1. Let the parameter $\theta$ be the midpoint of the range. For a sample of $n$ from a distribution of this family, we look for a sufficient statistic.

The parameter $\theta$ is a parameter of selection. Hence the statistic of selection $T$ is f-sufficient. $T$ is the $n$th power of the interval obtained from the intersection of all the unit intervals which contain the $n$ points. It is equivalent to the closed interval $[x_L, x_G]$ where $x_L$ and $x_G$ are respectively the least and greatest values in the sample. Thus $x_L$ and $x_G$ are jointly f-sufficient for $\theta$. It is easy to verify that the conditions of Lemma 4 are fulfilled and hence $x_L$ and $x_G$ are jointly sufficient for $\theta$.

EXAMPLE 2. To a Poisson variable with fixed mean is added an "error" uniformly distributed between $-\mu$ and $+\mu$ where $\mu < \frac{1}{2}$. Let $\mu$ be the unknown parameter. For a sample of $n$ from this distribution, we look for a sufficient statistic.

This class of distributions satisfies the hypothesis of Theorem 3. Therefore a real valued sufficient statistic $T$ exists. $T$ is equal to the largest deviation for the $n$ sample values: the deviation for $x_i$ is defined to be the absolute value of the distance from $x_i$ to the nearest integer.

If we let $(x_i)$ be the integer nearest the sample element $x_i$, then

$$T = \sup_{i=1,\cdots,n} | x_i - (x_i) | .$$

EXAMPLE 3. To a Poisson variable with parameter $m$ is added an "error" uniformly distributed between $-\mu$ and $+\mu$, where $\mu < \frac{1}{2}$. Let $(m, \mu)$ be the unknown parameter. For a sample of $n$, we find the minimal f-sufficient statistic and then check for sufficiency.

The statistic of selection $T$ is as calculated in Example 2. (We only consider points in $R^1$ for which the coordinates are all $> -\frac{1}{2}$). The carrier of the distribution $S_\mu$ can be defined by

$$S_\mu = [\{x \mid | x - (x)| < \mu\}]^n.$$

Therefore

$$f_{\mu,m}(\bar{x}) = \phi_{S_\mu}(\bar{x}) \cdot n(\mu) \prod_{i=1}^{n} \frac{m^{(x_i)}}{(x_i)!} e^{-m}.$$

We wish to define $\tilde{f}_{\mu,m}(\bar{x})$ to facilitate the calculation of the minimal f-sufficient statistic and yet consistent with the functional form prescribed in Theorem 1. The obvious definition fulfills the necessary conditions:

$$\tilde{f}_{\mu,m}(\bar{x}) = \prod_{i=1}^{n} \frac{m^{(x_i)}}{(x_i)!} e^{-m}.$$

The minimal f-sufficient statistic for $\{\tilde{f}_{\mu,m}(x)\}$ is $\sum(x_i)$.

By Theorem 1 the minimal f-sufficient statistic for $\{f_{\mu,m}(x)\}$ is

$$\left[ \sup_{i=1,\cdots,n} | x_i - (x_i) |, \quad \sum_{i=1}^{n} (x_i) \right].$$

Since the conditions of Lemma 4 are easily seen to be fulfilled, the above statistic is also sufficient.

## REFERENCES

[1] B. O. KOOPMAN, "On distributions admitting a sufficient statistic," *Trans. Am. Math. Soc.*, Vol. 39 (1936), pp. 399–409.

[2] PAUL R. HALMOS AND L. J. SAVAGE, "Applications of the Radon-Nikodym theorem to the theory of sufficient statistics," *Annals of Math. Stat.*, Vol. 20 (1949), pp. 225–241.

[3] R. C. DAVIS, "On minimum variance in nonregular estimation," *Annals of Math. Stat.*, Vol. 22 (1951), pp. 43–57.

[4] E. L. LEHMANN AND HENRY SCHEFFÉ, "Completeness, similar regions, and unbiased estimation," *Sankhyā*, Vol. 10 (1950), pp. 305–340.

# ORTHOGONAL ARRAYS OF INDEX UNITY[1]

By K. A. Bush

*University of North Carolina and University of Illinois*

**Summary.** In this paper we shall proceed to generalize the notion of a set of orthogonal Latin squares, and we term this extension an orthogonal array of index unity. In Section 2 we secure bounds for the number of constraints which are the counterpart of the familiar theorem which states that the number of mutually orthogonal Latin squares of side $s$ is bounded above by $s - 1$. Curiously, our bound depends upon whether $s$ is odd or even. In Section 3 we give a method of constructing these arrays by considering a class of polynomials with coefficients in the finite Galois field $GF(s)$, where $s$ is a prime or a power of a prime. In the concluding section we give a brief discussion of designs based on these arrays.

**1. Introduction.** Let a set of $s$ integers $0, 1, \cdots, s - 1$ be arranged in an $s \times s$ square in such a way that every integer occurs $s$ times. If each integer occurs once and only once in every row and column, the square is said to be a Latin square of side $s$. Two squares are said to be orthogonal to one another if, when one square is superimposed upon the other square, every number of the first occurs once and only once with every number of the second square. To the set of at most $s - 1$ Latin squares which are mutually orthogonal, we may adjoin two other squares which are not Latin squares but which are orthogonal to each other and to every other Latin square in the orthogonal set. The first of these squares is constructed by taking each element of the first row as 0, each element of the second row as 1, and so on. The second square is the transpose of the first square. Conversely it may be noted that any square orthogonal to these two squares must be a Latin square. Thus a total of $s + 1$ orthogonal squares is possible at best, and it is known that this bound is attainable when $s$ is a prime or a power of a prime [1]. When this bound is attained, we say that we have a complete set of orthogonal squares. As an example of a complete set, we might choose $s = 3$ and write

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | | 0 | 1 | 2 | | 0 | 1 | 2 | | 0 | 1 | 2 |
| 1 | 1 | 1 | | 0 | 1 | 2 | | 1 | 2 | 0 | | 2 | 0 | 1 |
| 2 | 2 | 2 | | 0 | 1 | 2 | | 2 | 0 | 1 | | 1 | 2 | 0 |

If we write in order the elements of each square in a line, we can display these squares in the following form:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 1 | 1 | 1 | 2 | 2 | 2 | [first square] |
| 0 | 1 | 2 | 0 | 1 | 2 | 0 | 1 | 2 | [second square] |

[1] This note is a revision of one part of the author's doctoral dissertation submitted to the University of North Carolina at Chapel Hill.

0 1 2 1 2 0 2 0 1 [third square]
0 1 2 2 0 1 1 2 0 [fourth square]

In this form we see that any two rows have the property that each one of the nine possible ordered pairs occurs exactly once when one row is superimposed on another row. We now generalize this basic idea.

Let us consider a matrix $A = [a_{ij}]$, where each $a_{ij}$ represents one of the integers $0, 1, \cdots, s-1$, $s > 1$. The matrix is rectangular with $N$ columns, which we shall call the blocks of the array, and $k$ rows. Consider all $t$-rowed submatrices of $N$ columns which can be formed from this array, $t \leqq k$. Each column of any $t$-rowed submatrix can be regarded as an ordered $t$-plet so that each $t$-rowed submatrix contains $N$ such $t$-plets. The matrix $A$ will be called an orthogonal array $[N, k, s, t]$ of *size* $N$, $k$ *constraints*, $s$ *levels*, *strength* $t$ and *index* $\lambda$ if each of the $C_t^k$ $t$-rowed $N$-columned submatrices that may be formed from the array contains every one of the $s^t$ possible ordered $t$-plets each repeated $\lambda$ times. It is clear that this definition implies that each row contains the $s$ integers $0, 1, \cdots, s-1$, each repeated $\lambda s^{t-1}$ times. We shall consider the case where $\lambda = 1$ and refer to such arrays as "orthogonal arrays of index unity." We shall consider arrays where $t > 2$ since the case $t = 2$ has been completely discussed and solved for $s$ a prime or a power of prime. References [1], [2], [7], [8], and [11] discuss this problem.

**2. Upper bounds for the number of constraints.** We shall use the notation $f(N, s, t)$ to denote the maximum number of constraints which are possible. Thus $f(N, s, 2) = s + 1$ when $s$ is a prime or a power of a prime. We prove first the

THEOREM. *If* $s \leqq t$, *then* $f(s^t, s, t) \leqq t + 1$.

PROOF. We shall make repeated use of the following fact: selecting any $t - 1$ rows, any specified $(t-1)$-plet must occur exactly $s$ times, for these $s$ identical $(t-1)$-plets may be enlarged to $t$-plets by the adjunction of another row in exactly $s$ ways since a given $t$-plet occurs once and only once. If possible let there exist $t + 2$ orthogonal rows. In order to facilitate notation we shall assume that one of the blocks consists entirely of zeros. This is justified because within each row we can always make a permutation of the elements $0, 1, 2, \cdots, s-1$ without destroying the orthogonality of the array. The general case proceeds along slightly different lines from the special cases $s = 2$ and $s = 3$ which we shall consider first.

*Proof for the case* $s = 2$. We divide the array into two parts. The first part, $A$, consists of three rows, and the second part, $B$, consists of $t - 1$ rows. There will be exactly two blocks which will have zero in every row of $B$. One of these blocks is our initial block consisting exclusively of zeros. In the second block the three rows of $A$ must contain the symbol 1, for otherwise there would be two identical $t$-plets. Now there will be two blocks which will have the symbol 1 in the first row of $B$ and zero in all other rows. Let these be the third and fourth

blocks. The first four blocks of the array we have just described will have the form:

| | | | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | . | . |
| $A$ | | 0 | 1 | . | . |
| | | 0 | 1 | . | . |
| | | 0 | 0 | 1 | 1 |
| $B$ | | 0 | 0 | 0 | 0 |
| | | . | . | . | . |
| | | 0 | 0 | 0 | 0 |

The third block already has $t - 2$ elements in common with each of the first two blocks. Therefore it can have at the most one more element in common with any of these blocks. Since the three rows of $A$ for the third block have to be filled in by using the symbols 0 and 1 only, we cannot avoid the situation in which it has at least two elements in part $A$ in common with one of the first two blocks. The theorem is thus proved for the case $s = 2$.

We now proceed to the case $s = 3$ using the same general type of argument although there are more complications. We divide the array into two parts as before, three rows in $A$ and $t - 1$ rows in $B$. In this case the first fifteen blocks of the array can be written as follows:

| | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | | | 0 | | | 0 | | | 0 | | | 0 | | | | | | | | | | | |
| $A$ | 0 | | | | 0 | | | 0 | | | 0 | | | | 0 | | | | | | | | | |
| | 0 | | | | | 0 | | | 0 | | | 0 | | | | 0 | | | | | | | | |
| | 0 | 0 | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | | | |
| $B$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 2 | 2 | 2 | | | | | | | | | |
| | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | | | | | | | | | |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | | | | | | | | |

The blank spaces in $A$ are to be filled in with 1 or 2. Obviously with each group of three blocks with identical $(t - 1)$-plets in $B$, we must include 0 from each row of $A$, and two 0's may not occur in the same block since then a $t$-plet identical with a $t$-plet from the first block would arise. We will encounter a somewhat similar type of arrangement in the general case shortly to be discussed. Each row of $B$ in the portion of the array exhibited contains only zeros with the exception of the first two rows.

Let us consider the four blocks in the array which have 0 in the top line of $A$ not counting the initial block. We now attempt to complete the second and third rows of $A$ for these four blocks. We note that these four blocks already have a common element in $A$ so that if two of the four blocks were completed using the same pair we would have two identical triplets in $A$. However any two blocks have at least $t - 3$ 0's in common in $B$ so that two identical $t$-plets would eventuate. Therefore the four pairs must be distinct. From the elements 1 and 2 we can make only four distinct ordered pairs so that each possible pair must be utilized in these four blocks.

In completing the second and third blocks we must therefore repeat these

pairs. Hence we can find two blocks with two elements in common in $A$. But this is impossible since (say) the second block has $t - 2$ 0's in common with every other block in the $B$ portion.

*Proof for the case* $s > 3$. We shall use the same notations and procedures employed in the proof of the previous cases. We shall have an initial block consisting of $t + 2$ zeros. We divide the array into an $A$ group with three rows and a $B$ group with $t - 1$ rows. We then write down the $s - 1$ blocks which contain only 0's in the rows of $B$. We shall call this group of blocks $F$. We follow this with a group of $s$ blocks which have the element 1 in the top row of $B$ and the element 0 in the remaining rows of $B$. This is followed by a group of $s$ blocks containing the element 2 in the top row of $B$, 0's elsewhere. Finally we arrive at a group of blocks which have the element $s - 1$ in the top row with 0's elsewhere. We have now secured $s - 1$ groups with $s$ blocks in a group in addition to the group $F$. We now generate $s - 1$ further groups each containing $s$ blocks by choosing the second row of $B$ as our non-zero row. Proceeding in this way we finally obtain $(t - 1)(s - 1)$ groups each containing $s$ blocks since there are $t - 1$ rows in $B$.

Consider one of these groups other than $F$ which have identical $(t - 1)$-plets in $B$. Over such a group the rows of $A$ must contain some permutation of the elements $0, 1, \cdots, s - 1$ in order that no $t$-plet be repeated. Hence the element 0 must occur once in each row of $A$ over such a group; further, it must not occur twice in the same block, for then a $t$-plet consisting exclusively of 0's would arise. There are thus $(s - 3)(t - 1)(s - 1)$ blocks in these groups (exclusive of $F$) which contain no element 0 in the three rows of $A$.

Over every row of $A$ each of the elements $1, 2, \cdots, s - 1$ occurs once in each group. If this condition is not met, then we either lose a $t$-plet or we secure more than one $t$-plet since in $B$ each group exhausts a given $(t - 1)$-plet. Let us denote by $e$ that element which occurs most frequently in the first row of $A$ in those blocks which do not contain the element 0 and are not in $F$. Then $e$ occurs in these blocks at least $(s - 3)(t - 1)$ times, for this is the number of blocks averaged over the number of elements, and there are $s - 1$ nonzero elements.

The element $e$ will also occur in some block of $F$ in the first row of $A$. Suppose that in this block the elements $a$ and $b$ occur in the second and third row. Now no element in the second row of $A$ can be the element $a$ if the same block contains the element $e$ in its first row, for then we would have two elements in common in $A$, and any block in $F$ has $t - 2$ elements in common with any other block in $B$; this would yield two identical $t$-plets. A similar remark applies to the element $b$ in the third row of $A$. Since we are considering only those blocks not in $F$ which contain no zero in any row of $A$, we may therefore form but $(s - 2)^2$ pairs when the element $e$ occurs in such a block.

However, the pairs in these blocks must all be distinct. If they were not, then we would have three elements in common in $A$, and always there are $t - 3$ zeros in common in $B$ so that again two $t$-plets would be identical. Furthermore, none of these pairs in the second and third rows may coincide with a pair

in the second and third row of $F$ by a like argument, that is, two elements alike in $A$, $t - 2$ elements alike in $B$. There are $s - 2$ pairs which remain in $F$ since we must obviously exclude the pair $(a, b)$ for which we have made a stronger statement. Now there are at least $(s - 3)(t - 1)$ distinct pairs not in $F$ so that we must have

$$(s - 2)^2 \geqq (s - 2) + (s - 3)(t - 1),$$

for the total number of allowable pairs must not be less than the total number of distinct pairs required. Upon simplification we obtain

$$(3 - s)(t - s + 1) \geqq 0,$$

which is a contradiction. This proves our theorem for $s > 3$.

We now prove a special result for the case $t = 3$. In the proof we require the following definition.

DEFINITION. *Two blocks will be called disjoint if corresponding elements are different, that is, $a_{ij} \neq a_{ik}$ for each value of $i$.*

THEOREM. *When $s$ is odd, $f(s^3, s, 3) \leqq s + 1$.*

PROOF. Suppose that $s + 2$ orthogonal rows can be constructed. Now each element is repeated $s^2$ times in each row and each pair of elements occurs $s$ times in any two rows. Consider the elements in the first block, and let us select a particular element in this block. With the other elements in the block we can form $s + 1$ pairs where the particular element selected is used as one component of the pairs. Taking into account the fact that each of these pairs occurs once in the initial block and that we have also here used one replication of our selected element, we have to form pairs in other blocks containing the element selected $(s + 1)(s - 1)$ times. But this exhausts the replications of the element chosen. Consequently any two blocks have either two elements in common or are disjoint. From the $s + 2$ elements in the initial block we can form $(s + 2)(s + 1)/2$ pairs which are repeated $s - 1$ times. But

$$(s + 2)(s + 1)(s - 1)/2 < s^3 - 1, \quad s > 1,$$

so that disjoint blocks exist. Let us consider two disjoint blocks, and all other blocks $(s - 1)$ in number which contain the same pair in the first two rows as the second of these disjoint blocks. We can then represent the situation as

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $a_1$ | $b_1$ | $b_1$ | $b_1$ | . | $b_1$ | . | . | . |
| $a_2$ | $b_2$ | $b_2$ | $b_2$ | . | $b_2$ | . | . | . |
| $a_3$ | $b_3$ | | | | | | | . |
| $a_4$ | $b_4$ | | | | | | | . |
| . | . | . | | | | | | . |
| $a_{s+2}$ | $b_{s+2}$ | . | | | | | | . |

where $a_i \neq b_i$, $i = 1, 2, \cdots, s + 2$. Since every triplet occurs exactly once, then the last $s$ elements in the last $s$ rows of the first block must occur exactly once in these $s - 1$ blocks. But since two blocks have either none or two elements in common, we get a contradiction since $s$ is odd.

We are now in a position to state the general theorem.

THEOREM. *The maximum number of constraints is bounded above by* $s + t - 1$ *if* $t \leqq s$ *and* $s$ *is even. In case* $s$ *is odd, we may refine the inequality to* $s + t - 2$ *when* $3 \leqq t \leqq s$.

PROOF. This theorem is an immediate consequence of the preceding theorem, for the existence of an array with $k + 1$ constraints and strength $t$ implies the existence of an array with $k$ constraints of strength $t - 1$ since we may select a particular row in the first array and consider the $s^{t-1}$ blocks which have 0 as the element in that row. Evidently the $k$ rows which remain form an orthogonal array of strength $t - 1$, for each $(t - 1)$-plet occurs exactly one time; otherwise with 0 as a common element some $t$-plet would occur more than once in the first array. Hence increasing $t$ by one can increase the number of constraints at most by one.

The inequality given above represents a substantial improvement over a similar inequality due to Rao [8]. Rao's result states that the number of constraints $k$ is bounded above by

$$s^t - 1 \geqq C_1^k(s-1) + \cdots + C_u^k(s-1)^u \qquad \text{if } t = 2u,$$

$$s^t - 1 \geqq C_1^k(s-1) + \cdots + C_u^k(s-1)^u + C_u^{k-1}(s-1)^{u+1} \qquad \text{if } t = 2u+1.$$

When $t = 3$, the Rao bound is $s + 2$, so that we have reduced the bound by unity if $s$ is odd. When $t = 4$ and $s = 2$, the Rao bound is 5, agreeing with our result. When $s = 3$ and $t = 4$, the bound is 6 against our 5; when $s = 4$ and $t = 4$, the bound is 7 against our 5; when $s = 5$ and $t = 4$, we have 9 against 7. The differential tends to increase both with increasing $s$ and increasing $t$. Thus when $t = 6$ and $s = 11$, the Rao bound is 22 against our 15.

**3. Construction of orthogonal arrays of index unity.** We now prove the following theorem which, in conjunction with a corollary, provides a complete answer to the problem of construction of orthogonal arrays of index unity and strength three when $s$ is a prime or a power of a prime.

THEOREM. *If* $s = p^n$, *where* $p$ *is a prime and* $s > t$, *then we can construct the array* $(s^t, s + 1, s, t)$.

PROOF. Let $GF(s)$ denote the finite Galois field with $s = p^n$ elements which may be denoted by $e_i$, $i = 0, 1, \cdots, s - 1$. Consider the polynomials

$$y_j(x) = a_{t-1}x^{t-1} + a_{t-2}x^{t-2} + \cdots + a_1x + a_0,$$

where the coefficients range over the field $GF(s)$. It is clear that there are $s^t$ such polynomials, since each of the $t$ coefficients can assume $s$ different values in the field, and the subscript $j$ ranges over the values $0, 1, \cdots, s^t - 1$. We now form an $s$ by $s^t$ array numbering the rows from 0 to $s - 1$ and the columns from 0 to $s^t - 1$. In the $i$th row and $j$th column we agree to insert the integer $u$ with the property

$$y_j(e_i) = e_u.$$

We assert that the resulting array is orthogonal of strength $t$ and index unity.

Suppose on the contrary that we can select $t$ rows such that two $t$-plets are identical, and let the associated polynomials be

$$(1) \qquad \begin{aligned} a_{t-1}x^{t-1} + a_{t-2}x^{t-2} + \cdots + a_1x + a_0 &= y_j(x), \\ a'_{t-1}x^{t-1} + a'_{t-2}x^{t-2} + \cdots + a'_1x + a'_0 &= y_{j'}(x). \end{aligned}$$

Let the rows be those rows which correspond to the Galois field elements $e_{i_1}, e_{i_2}, \cdots, e_{i_t}$. Upon inserting the element $e_{i_r}$ in equations (1) above and subtracting these two equations, we have

$$A_{t-1}e_{i_r}^{t-1} + A_{t-2}e_{i_r}^{t-2} + \cdots + A_1e_{i_r} + A_0 = 0,$$

where $A_u = a'_u - a_u$. As we allow $r$ to range from 1 to $t$, we secure $t$ homogeneous equations linear in the $t$ quantities $A_{t-1}, A_{t-2}, \cdots, A_1, A_0$. Not all the $A$'s can be zero, since then the two polynomials would be identical. Consequently for a solution to exist the determinant of the matrix $V$, where

$$V = \begin{bmatrix} e_{i_1}^{t-1} & e_{i_1}^{t-2} & \cdots & e_{i_1} & 1 \\ e_{i_2}^{t-1} & e_{i_2}^{t-2} & \cdots & e_{i_2} & 1 \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ e_{i_t}^{t-1} & e_{i_t}^{t-2} & \cdots & e_{i_t} & 1 \end{bmatrix},$$

must vanish. However, it is well known that this matrix of Vandermonde type has the property

$$\det [V] = \prod_{u<v} (e_{i_u} - e_{i_v}),$$

and so is nonsingular unless $e_{i_u} = e_{i_v}$, which is clearly impossible since the rows were all distinct.

We have thus shown that all $t$-plets are distinct, and this completes the demonstration that we can construct $s$ rows. However, we can adjoin one more row in an orthogonal manner by assigning the value $u$ to those columns which are associated with the polynomial whose leading coefficient is $e_u$. In this case $A_{t-1} = 0$, and we have $t - 1$ equations in $t - 1$ unknowns. Here for a solution to exist we must demand that the Vandermonde determinant of order $t - 2$ vanish, and this is impossible. (We use *order* to indicate the highest degree of terms in the matrix.)

For the special case $t = 3$, it is possible to add yet another orthogonal row when $s = 2^n$ by assigning the value $u$ to those columns which have $e_u$ as the coefficient of the term of degree $t - 2$. It is clear that any $t$-plet which is constructed by using two components from these two rows and $t - 2$ components from the original $s$ rows occurs just once. Here both $A_{t-1}$ and $A_{t-2}$ vanish, and we have a nonvanishing determinant of Vandermonde type of order $t - 3$.

On the other hand, if our $t$-plet includes only the final row, the resultant matrix is no longer of Vandermonde type and may be singular. This cannot

happen over fields of characteristic 2 when $t = 3$, for we obtain the matrix:

$$\begin{bmatrix} e_{i_1}^2 & 1 \\ e_{i_2}^2 & 1 \end{bmatrix}.$$

The determinant of the matrix is $(e_{i_1} - e_{i_2})(e_{i_1} + e_{i_2})$, and since every element is its own additive inverse in fields of characteristic 2, it is clear that the matrix is nonsingular. We may therefore state the

COROLLARY. *For the array* $(s^3, k, s = p^n, 3)$, *where* $p$ *is a prime, we have*

$$f(s^3, s, 3) = s + 1, s = p^n, \qquad p \text{ an odd prime},$$
$$f(s^3, s, 3) = s + 2, s = 2^n,$$

by virtue of this construction and the inequalities of Section 2.

THEOREM. *If* $s \leqq t$, $f(s^t, s, t) = t + 1$, *where* $s = p^n$, $p$ *a prime.*

PROOF. In Section 2, we showed that for this case $f(s^t, s, t) \leqq t + 1$. We now show that this bound is attainable by using the same constructive procedure employed in securing the additional rows in the last case we discussed. For example, the third row will have the integer $u$ in every column which has $e_u$ as the coefficient of the term of degree $t - 3$. In this way we clearly succeed in constructing an array of strength $t$ and index unity which is orthogonal. To this array we can add another row by the process we first used where we now substitute into the various polynomials one of the nonzero elements of the field. For convenience we may as well associate the additional row with the unity of our field. It is clear that the matrix is nonsingular; indeed, it consists of a single nonzero term.

We may summarize these results when $t > 3$ as follows:

(i) If $s = 2^n > t$, then $s + 1 \leqq f(s^t, s, t) \leqq s + t - 1$.

(ii) If $s = p^n > t$, where $p$ is an odd prime, then $s + 1 \leqq f(s^t, s, t) \leqq s + t - 2$.

**4. Uses of orthogonal arrays.** A problem which often arises in the design of an experiment is that of ascertaining the effect of quantitative or qualitative alterations in the various components upon some measurable characteristic of the complete assembly. The traditional case is that of $k$ fertilizers each of which can be applied at $s$ different levels. To carry out a complete factorial experiment would require $s^k$ plots or *assemblies* to use a more general term applicable to any type of experiment, agronomic or otherwise.

Rao [8] has shown that if the design is an array of strength $t$, then the estimates of main effects and interactions are unaffected by interactions of order greater than one and less than $t - 1$, but the estimate of error is enhanced by their presence. If $t$ is even, we can measure interactions up to order $t/2$ supposing the higher order interactions absent. When $t$ is odd, we can measure interactions up to order $(t - 1)/2$. Thus an array of appropriate strength must be chosen to handle those interactions which are deemed important, but it will not in general be necessary to construct an array of size $s^k$.

## REFERENCES

[1] R. C. Bose, "On the application of the properties of Galois fields to the construction of hyper Graeco-Latin squares," *Sankhyā*, Vol. 3 (1938), pp. 328–338.

[2] R. C. Bose and K. R. Nair, "On complete sets of Latin squares," *Sankhyā*, Vol. 5 (1941), pp. 361–382.

[3] R. H. Bruck and H. J. Ryser, "The nonexistence of certain finite projective planes," *Canadian Jour. Math.*, Vol. 1 (1949), pp. 88–94.

[4] R. A. Fisher, "A system of confounding for factors with more than two alternatives giving completely orthogonal cubes and higher powers," *Annals of Eugenics*, Vol. 12 (1942), pp. 283–290.

[5] R. A. Fisher and F. Yates, "The 6 × 6 Latin squares," *Proc. Cambridge Philos. Soc.*, Vol. 30 (1934), pp. 482–507.

[6] H. F. MacNeish, "Euler's squares," *Annals of Mathematics*, Vol. 23 (1922), pp. 221–227.

[7] H. B. Mann, "The construction of orthogonal Latin squares," *Annals of Math. Stat.*, Vol. 13 (1942), pp. 418–423.

[8] C. R. Rao, "Factorial arrangements derivable from combinatorial arrangements of arrays," *Jour. Roy. Stat. Soc.*, Vol. 9 (1947), pp. 128–139.

[9] C. R. Rao, "On a class of arrangements," *Proc. Edinburgh Math. Soc.*, Vol. 8 (1949), pp. 119–125.

[10] W. L. Stevens, "The completely orthogonalized Latin square," *Annals of Eugenics*, Vol. 9 (1939), pp. 82–93.

[11] G. Tarry, "Le problème des 36 officiers," *Comptes Rendus de l'Association Française pour l'Avancement de Science Naturel*, Vol. 1 (1900), pp. 122–123; Vol. 2 (1901), pp. 170–203.

# DISTRIBUTION TABLE FOR THE DEVIATION BETWEEN TWO SAMPLE CUMULATIVES[1]

By Frank J. Massey, Jr.

*University of Oregon*

Suppose $x_1 < x_2 < \cdots < x_{N_1}$, $y_1 < y_2 < \cdots < y_{N_2}$ are the ordered results of two random samples, of sizes $N_1$ and $N_2$, taken from two populations with cumulative distribution functions $F_1(x)$ and $F_2(y)$.

Let $d = \max | S_{N_1}(x) - S_{N_2}(x) |$, where $S_{N_1}(x)$ is the proportion of the $N_1$ observed values $x_i$ less than or equal to $x$, and $S_{N_2}(x)$ is the proportion of the $N_2$ observed values $y_i$ less than or equal to $x$. If $F_1(x) \equiv F_2(x) \equiv F(x)$ and if $F(x)$ is continuous then the distribution of $d$ is independent of $F(x)$. Table 1 lists, for $N_1, N_2 \leqq 10$, all possible values of $d_\alpha$ and $\alpha$ such that

$$P\{d \leqq d_\alpha\} = \alpha.$$

Table 2 lists certain other selected values of $N_1$, $N_2$, and $d_\alpha$.

The method used in finding the values is that described in [1].

## REFERENCE

[1] Frank J. Massey, Jr., "The distribution of the maximum deviation between two sample cumulative step functions," *Annals of Math. Stat.*, Vol. 22 (1951), pp. 125–128.

[1] The method of computing the values tabled was developed under contract N6-onr-218/IV with the Office of Naval Research (see [1]). The main part of the table was made possible by a grant from the Graduate School of the University of Oregon.

TABLE 1

Values of $\alpha$ and $d_\alpha$ such that for $N_1, N_2 \leqq 10$, $P\{d \leqq d_\alpha\} = \alpha$. Headings are in the form $(N_1, N_2, d_\alpha)$ with the $d_\alpha$ expressed as a fraction with the least common multiple of $N_1$ and $N_2$ in the denominator. All possible values of $d_\alpha$ are listed in the table.

| (1, 1, h/1) | | (1, 2, h/2) | | (1, 3, h/3) | | (1, 4, h/4) | | (1, 5, h/5) | | (1, 6, h/6) | | (1, 7, h/7) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| h | α | h | α | h | α | h | α | h | α | h | α | h | α |
| 1 | 1 | 1 | .33 | 2 | .5 | 2 | .2 | 3 | .33333 | 3 | .14286 | 4 | .25 |
| | | 2 | 1 | 3 | 1 | 3 | .6 | 4 | .66667 | 4 | .42857 | 5 | .50 |
| | | | | | | 4 | 1 | 5 | 1 | 5 | .71429 | 6 | .75 |
| | | | | | | | | | | 6 | 1 | 7 | 1 |

| (1, 8, h/8) | | (1, 9, h/9) | | (1, 10, h/10) | | | | (2, 2, h/2) | | (2, 3, h/6) | | (2, 4, h/4) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| h | α | h | α | h | α | | | h | α | h | α | h | α |
| 4 | .11111 | 5 | .2 | 5 | .09091 | | | 1 | .66667 | 2 | .1 | 1 | .06667 |
| 5 | .33333 | 6 | .4 | 6 | .27273 | | | 2 | 1 | 3 | .4 | 2 | .60000 |
| 6 | .55556 | 7 | .6 | 7 | .45455 | | | | | 4 | .8 | 3 | .86667 |
| 7 | .77778 | 8 | .8 | 8 | .63636 | | | | | 6 | 1 | 4 | 1 |
| 8 | 1 | 9 | 1 | 9 | .81818 | | | | | | | | |
| | | | | 10 | 1 | | | | | | | | |

| (2, 5, h/10) | | (2, 6, h/6) | | (2, 7, h/14) | | (2, 8, h/8) | | (2, 9, h/18) | | (2, 10, h/10) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| h | α | h | α | h | α | h | α | h | α | h | α | |
| 3 | .04762 | 2 | .14286 | 4 | .02778 | 2 | .02222 | 5 | .01818 | 3 | .06061 | |
| 4 | .19048 | 3 | .57143 | 5 | .11111 | 3 | .20000 | 6 | .07273 | 4 | .24242 | |
| 5 | .42857 | 4 | .78571 | 6 | .25000 | 4 | .55556 | 7 | .16364 | 5 | .54545 | |
| 6 | .71429 | 5 | .92857 | 7 | .44444 | 5 | .73333 | 8 | .29091 | 6 | .69697 | |
| 8 | .90476 | 6 | 1 | 8 | .66667 | 6 | .86667 | 9 | .45455 | 7 | .81818 | |
| 10 | 1 | | | 10 | .83333 | 7 | .95556 | 10 | .63636 | 8 | .90909 | |
| | | | | 12 | .94444 | 8 | 1 | 12 | .78182 | 9 | .96970 | |
| | | | | 14 | 1 | | | 14 | .89091 | 10 | 1 | |
| | | | | | | | | 16 | .96364 | | | |
| | | | | | | | | 18 | 1 | | | |

| (3, 3, h/3) | | (3, 4, h/12) | | (3, 5, h/15) | | (3, 6, h/6) | | (3, 7, h/21) | | (3, 8, h/24) | | (3, 9, h/9) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| h | α | h | α | h | α | h | α | h | α | h | α | h | α |
| 1 | .4 | 3 | .02857 | 4 | .03571 | 1 | .01191 | 5 | .00833 | 5 | .00606 | 2 | .03636 |
| 2 | .9 | 4 | .11429 | 5 | .14286 | 2 | .32143 | 6 | .06667 | 6 | .02424 | 3 | .29091 |
| 3 | 1 | 5 | .34286 | 6 | .32143 | 3 | .66667 | 7 | .15000 | 7 | .07273 | 4 | .54545 |
| | | 6 | .60000 | 7 | .53571 | 4 | .90476 | 8 | .30000 | 8 | .16364 | 5 | .76364 |
| | | 8 | .77143 | 9 | .71429 | 5 | .97619 | 9 | .46667 | 9 | .29091 | 6 | .90909 |
| | | 9 | .94286 | 10 | .85714 | 6 | 1 | 11 | .60000 | 10 | .43636 | 7 | .96364 |
| | | 12 | 1 | 12 | .96429 | | | 12 | .75000 | 12 | .57576 | 8 | .99091 |
| | | | | 15 | 1 | | | 14 | .83333 | 13 | .69697 | 9 | 1 |
| | | | | | | | | 15 | .93333 | 15 | .80606 | | |
| | | | | | | | | 18 | .98333 | 16 | .87879 | | |
| | | | | | | | | 21 | 1 | 18 | .95152 | | |
| | | | | | | | | | | 21 | .98788 | | |
| | | | | | | | | | | 24 | 1 | | |

TABLE 1—*Continued*

| (3, 10, $h$/30) | | (3, 10, $h$/30) (*cont.*) | | (4, 4, $h$/4) | | (4, 5, $h$/20) | | (4, 6, $h$/12) | | (4, 7, $h$/28) | | (4, 7, $h$/28) (*cont.*) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ |
| 6 | .00350 | 18 | .81119 | 1 | .22857 | 5 | .03175 | 2 | .00476 | 5 | .00303 | 17 | .87879 |
| 7 | .01399 | 20 | .83566 | 2 | .77143 | 6 | .12698 | 3 | .07619 | 6 | .01212 | 20 | .93333 |
| 8 | .04196 | 21 | .93007 | 3 | .97143 | 7 | .25397 | 4 | .30476 | 7 | .04848 | 21 | .96970 |
| 9 | .09441 | 24 | .97203 | 4 | 1 | 8 | .43651 | 5 | .44762 | 8 | .10909 | 24 | .99394 |
| 10 | .16783 | 27 | .99301 | | | 10 | .57143 | 6 | .70476 | 9 | .21818 | 28 | 1 |
| 11 | .27972 | 30 | 1 | | | 11 | .71429 | 7 | .81905 | 10 | .33939 | | |
| 12 | .40210 | | | | | 12 | .85714 | 8 | .90476 | 12 | .46364 | | |
| 14 | .50699 | | | | | 15 | .92063 | 9 | .95238 | 13 | .58485 | | |
| 15 | .62937 | | | | | 16 | .98413 | 10 | .99048 | 14 | .69091 | | |
| 17 | .71329 | | | | | 20 | 1 | 12 | 1 | 16 | .78788 | | |

| (4, 8, $h$/8) | | (4, 9, $h$/36) | | (4, 10, $h$/20) | | | | (5, 5, $h$/5) | | (5, 6, $h$/30) | | (5, 7, $h$/35) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | | | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ |
| 1 | .00202 | 6 | .00140 | 3 | .00100 | | | 1 | .12698 | 5 | .00216 | 6 | .00253 |
| 2 | .16364 | 7 | .00559 | 4 | .01518 | | | 2 | .64286 | 6 | .00866 | 7 | .01010 |
| 3 | .48687 | 8 | .02238 | 5 | .08092 | | | 3 | .92064 | 7 | .03463 | 8 | .04040 |
| 4 | .77778 | 9 | .05035 | 6 | .22478 | | | 4 | .99206 | 8 | .10390 | 9 | .09091 |
| 5 | .91515 | 10 | .11329 | 7 | .31269 | | | 5 | 1 | 9 | .18182 | 10 | .16162 |
| 6 | .97980 | 11 | .18881 | 8 | .50050 | | | | | 10 | .31169 | 11 | .26263 |
| 7 | .99596 | 12 | .29231 | 9 | .59041 | | | | | 12 | .40909 | 13 | .36111 |
| 8 | 1 | 14 | .38182 | 10 | .74026 | | | | | 13 | .52597 | 14 | .45455 |
| | | 15 | .48252 | 11 | .81219 | | | | | 14 | .64286 | 15 | .56566 |
| | | 16 | .59580 | 12 | .87413 | | | | | 15 | .76190 | 16 | .67172 |
| | | 18 | .66434 | 13 | .91608 | | | | | 18 | .82251 | 18 | .76263 |
| | | 19 | .74825 | 14 | .95405 | | | | | 19 | .89177 | 20 | .83333 |
| | | 20 | .83497 | 15 | .97003 | | | | | 20 | .95238 | 21 | .88384 |
| | | 23 | .88531 | 16 | .99001 | | | | | 24 | .97403 | 23 | .93434 |
| | | 24 | .93846 | 18 | .99800 | | | | | 25 | .99567 | 25 | .96970 |
| | | 27 | .95804 | 20 | 1 | | | | | 30 | 1 | 29 | .98485 |
| | | 28 | .98601 | | | | | | | | | 30 | .99747 |
| | | 32 | .99720 | | | | | | | | | 35 | 1 |
| | | 36 | 1 | | | | | | | | | | |

| (5, 8, $h$/40) | | (5, 8, $h$/40) (*cont.*) | | (5, 9, $h$/45) | | (5, 9, $h$/45) (*cont.*) | | (5, 10, $h$/10) | | | (6, 6, $h$/6) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | | $h$ | $\alpha$ |
| 6 | .00078 | 20 | .76535 | 7 | .00100 | 21 | .69930 | 1 | .00033 | | 1 | .06926 |
| 7 | .00311 | 22 | .83528 | 8 | .00400 | 22 | .77423 | 2 | .08092 | | 2 | .52597 |
| 8 | .01243 | 24 | .87413 | 9 | .01598 | 25 | .82717 | 3 | .34532 | | 3 | .85714 |
| 9 | .03730 | 25 | .92075 | 10 | .03596 | 26 | .88112 | 4 | .64935 | | 4 | .97403 |
| 10 | .08392 | 27 | .95804 | 11 | .08092 | 27 | .91409 | 5 | .83417 | | 5 | .99784 |
| 11 | .14918 | 29 | .97980 | 12 | .13487 | 30 | .94406 | 6 | .93939 | | 6 | 1 |
| 12 | .23077 | 32 | .99068 | 13 | .20879 | 31 | .97203 | 7 | .98069 | | | |
| 14 | .31624 | 35 | .99845 | 15 | .28571 | 35 | .98601 | 8 | .99600 | | | |
| 15 | .41414 | 40 | 1 | 16 | .37363 | 36 | .99401 | 9 | .99933 | | | |
| 16 | .50039 | | | 17 | .45854 | 40 | .99900 | 10 | 1 | | | |
| 17 | .59984 | | | 19 | .53946 | 45 | 1 | | | | | |
| 19 | .68376 | | | 20 | .61938 | | | | | | | |

TABLE 1—*Continued*

| (6, 7, $h$/42) | | (6, 7, $h$/42) (*cont.*) | | (6, 8, $h$/24) | | (6, 8, $h$/24) (*cont.*) | | (6, 9, $h$/18) | | (6, 10, $h$/30) | | (6, 10, $h$/30) (*cont.*) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ |
| 6 | .00058 | 28 | .93240 | 3 | .00033 | 20 | .99467 | 2 | .00020 | 4 | .00050 | 20 | .96853 |
| 7 | .00233 | 29 | .96154 | 4 | .00533 | 21 | .99933 | 3 | .01279 | 5 | .00799 | 21 | .98102 |
| 8 | .00932 | 30 | .98485 | 5 | .04795 | 24 | 1 | 4 | .10230 | 6 | .04046 | 22 | .99101 |
| 9 | .03730 | 35 | .99126 | 6 | .14685 | | | 5 | .22058 | 7 | .11239 | 24 | .99600 |
| 10 | .07459 | 36 | .99883 | 7 | .21612 | | | 6 | .43716 | 8 | .16459 | 25 | .99825 |
| 11 | .12821 | 42 | 1 | 8 | .36264 | | | 7 | .59161 | 9 | .28746 | 27 | .99975 |
| 12 | .21970 | | | 9 | .53979 | | | 8 | .72647 | 10 | .42258 | 30 | 1 |
| 14 | .28846 | | | 10 | .62371 | | | 9 | .82418 | 11 | .49451 | | |
| 15 | .37762 | | | 11 | .69830 | | | 10 | .90529 | 12 | .62887 | | |
| 16 | .47203 | | | 12 | .80753 | | | 11 | .93886 | 13 | .69505 | | |
| 17 | .56643 | | | 13 | .86081 | | | 12 | .97203 | 14 | .75125 | | |
| 18 | .66725 | | | 14 | .90743 | | | 13 | .98601 | 15 | .82992 | | |
| 21 | .72261 | | | 15 | .93939 | | | 14 | .99401 | 16 | .87488 | | |
| 22 | .78788 | | | 16 | .95738 | | | 15 | .99720 | 17 | .90784 | | |
| 23 | .85315 | | | 17 | .97736 | | | 16 | .99960 | 18 | .93357 | | |
| 24 | .90909 | | | 18 | .99068 | | | 18 | 1 | 19 | .95804 | | |

| (7, 7, $h$/7) | | (7, 8, $h$/56) | | (7, 8, $h$/56) (*cont.*) | | (7, 9, $h$/63) | | (7, 9, $h$/63) (*cont.*) | | (7, 10, $h$/70) | | (7, 10, $h$/70) (*cont.*) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ |
| 1 | .03730 | 7 | .00016 | 28 | .85859 | 8 | .00017 | 29 | .78234 | 8 | .00005 | 29 | .68511 |
| 2 | .42483 | 8 | .00062 | 32 | .88190 | 9 | .00070 | 31 | .83129 | 9 | .00021 | 30 | .71858 |
| 3 | .78788 | 9 | .00249 | 33 | .91298 | 10 | .00280 | 33 | .87325 | 10 | .00082 | 32 | .76743 |
| 4 | .94697 | 10 | .00995 | 34 | .93629 | 11 | .01119 | 35 | .90210 | 11 | .00329 | 33 | .81062 |
| 5 | .99184 | 11 | .02984 | 35 | .96737 | 12 | .02517 | 36 | .92133 | 12 | .00987 | 35 | .84461 |
| 6 | .99942 | 12 | .05221 | 40 | .97576 | 13 | .04476 | 38 | .94493 | 13 | .02221 | 36 | .88348 |
| 7 | 1 | 13 | .08951 | 41 | .98695 | 14 | .07710 | 40 | .96591 | 14 | .03949 | 39 | .91310 |
| | | 14 | .15338 | 42 | .99534 | 15 | .12483 | 42 | .97902 | 15 | .06911 | 40 | .92842 |
| | | 16 | .18943 | 48 | .99751 | 17 | .17238 | 45 | .98514 | 16 | .10587 | 42 | .94591 |
| | | 17 | .26434 | 49 | .99969 | 18 | .22500 | 47 | .99248 | 18 | .14937 | 43 | .96390 |
| | | 18 | .33986 | 56 | 1 | 19 | .28636 | 49 | .99720 | 19 | .19616 | 46 | .97830 |
| | | 19 | .41538 | | | 20 | .35245 | 55 | .99860 | 20 | .24666 | 49 | .98416 |
| | | 20 | .49231 | | | 21 | .42727 | 56 | .99983 | 21 | .30769 | 50 | .98900 |
| | | 21 | .57824 | | | 22 | .50122 | 63 | 1 | 22 | .37392 | 53 | .99393 |
| | | 24 | .62720 | | | 23 | .56976 | | | 23 | .43979 | 56 | .99825 |
| | | 25 | .65703 | | | 26 | .63199 | | | 25 | .50273 | 61 | .99918 |
| | | 26 | .71795 | | | 27 | .67622 | | | 26 | .56659 | 63 | .99990 |
| | | 27 | .77762 | | | 28 | .72990 | | | 28 | .62176 | 70 | 1 |

TABLE 1—*Concluded*

| (8, 8, $h$/8) | | (8, 9, $h$/72) | | (8, 9, $h$/72) (*cont.*) | | (8, 9, $h$/72) (*cont.*) | | (8, 10, $h$/40) | | (8, 10, $h$/40) (*cont.*) | | (8, 10, $h$/40) (*cont.*) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ |
| 1 | .01989 | 8 | .00004 | 22 | .35985 | 40 | .94406 | 5 | .00037 | 18 | .79633 | 32 | .99918 |
| 2 | .33986 | 9 | .00016 | 23 | .42390 | 45 | .95311 | 6 | .00585 | 19 | .83404 | 35 | .99959 |
| 3 | .71733 | 10 | .00066 | 24 | .49720 | 46 | .96643 | 7 | .02340 | 20 | .87975 | 36 | .99995 |
| 4 | .91298 | 11 | .00263 | 27 | .53970 | 47 | .97976 | 8 | .06913 | 21 | .90123 | 40 | 1 |
| 5 | .98135 | 12 | .01053 | 28 | .59301 | 48 | .98881 | 9 | .10176 | 22 | .92317 | | |
| 6 | .99751 | 13 | .02106 | 29 | .64829 | 54 | .99169 | 10 | .17729 | 23 | .95014 | | |
| 7 | .99985 | 14 | .03620 | 30 | .70358 | 55 | .99572 | 11 | .27725 | 24 | .96270 | | |
| 8 | 1 | 15 | .06203 | 31 | .75755 | 56 | .99860 | 12 | .39760 | 25 | .96974 | | |
| | | 16 | .10629 | 32 | .80424 | 63 | .99926 | 13 | .45775 | 26 | .97934 | | |
| | | 18 | .13957 | 36 | .82682 | 64 | .99992 | 14 | .51469 | 27 | .98798 | | |
| | | 19 | .18326 | 37 | .85767 | 72 | 1 | 15 | .60825 | 28 | .99301 | | |
| | | 20 | .23990 | 38 | .89058 | | | 16 | .67045 | 30 | .99502 | | |
| | | 21 | .29988 | 39 | .92143 | | | 17 | .75246 | 31 | .99758 | | |

| (9, 9, $h$/9) | | (9, 10, $h$/90) | | (9, 10, $h$/90) (*cont.*) | | (9, 10, $h$/90) (*cont.*) | | (9, 10, $h$/90) (*cont.*) | | (9, 10, $h$/90) (*cont.*) | | (10, 10, $h$/10) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ |
| 1 | .01054 | 9 | .00001 | 20 | .09616 | 32 | .55754 | 45 | .91608 | 70 | .99725 | 1 | .00554 |
| 2 | .26989 | 10 | .00004 | 21 | .12626 | 33 | .60742 | 50 | .92537 | 71 | .99864 | 2 | .21307 |
| 3 | .64829 | 11 | .00017 | 22 | .16572 | 34 | .65713 | 51 | .93966 | 72 | .99959 | 3 | .58248 |
| 4 | .87413 | 12 | .00069 | 23 | .21307 | 35 | .70580 | 52 | .95544 | 80 | .99978 | 4 | .83218 |
| 5 | .96643 | 13 | .00277 | 24 | .26042 | 36 | .74847 | 53 | .96564 | 81 | .99998 | 5 | .94755 |
| 6 | .99371 | 14 | .00831 | 25 | .30864 | 40 | .76993 | 54 | .97902 | 90 | 1 | 6 | .98766 |
| 7 | .99926 | 15 | .01455 | 26 | .35336 | 41 | .79970 | 60 | .98236 | | | 7 | .99794 |
| 8 | .99996 | 16 | .02494 | 28 | .42493 | 42 | .83218 | 61 | .98766 | | | 8 | .99978 |
| 9 | 1 | 17 | .04274 | 30 | .46141 | 43 | .86465 | 62 | .99281 | | | 9 | .99999 |
| | | 18 | .07323 | 31 | .50783 | 44 | .89442 | 63 | .99630 | | | 10 | 1 |

TABLE 2

Values of $\alpha$ and $d_\alpha$ such that $P\{d \leq d_\alpha\} = \alpha$ for certain values of $N_1$, $N_2$ and $d_\alpha$. Headings are expressed as in Table 1 in the form $(N_1, N_2, d_\alpha)$.

| (3, 12, $h$/12) | | (4, 12, $h$/12) | | (4, 16, $h$/16) | | (5, 15, $h$/15) | | (5, 20, $h$/20) | | (6, 12, $h$/12) | | (6, 18, $h$/18) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ |
| 2 | .00220 | 5 | .59560 | 2 | .00021 | 6 | .65738 | 2 | .00002 | 1 | .00005 | 7 | .69767 |
| 3 | .05934 | 6 | .78462 | 3 | .01672 | 7 | .79463 | 3 | .00457 | 2 | .03927 | 8 | .81794 |
| 4 | .27473 | 7 | .88791 | 4 | .12900 | 8 | .89048 | 4 | .05882 | 3 | .24036 | 9 | .89899 |
| 5 | .47473 | 8 | .95165 | 5 | .29164 | 9 | .94814 | 5 | .17397 | 4 | .53200 | 10 | .94655 |
| 6 | .66154 | 9 | .98352 | 6 | .47864 | 10 | .97665 | 6 | .33618 | 5 | .74618 | 11 | .97459 |
| 7 | .81099 | 10 | .99451 | 7 | .65243 | 11 | .99123 | 7 | .50941 | 6 | .88537 | 12 | .98908 |
| 8 | .91209 | 11 | .99890 | 8 | .78947 | 12 | .99729 | 8 | .66319 | 7 | .95367 | 13 | .99563 |
| 9 | .95604 | | | 9 | .87203 | 13 | .99923 | 9 | .77216 | 8 | .98513 | | |
| 10 | .98242 | | | 10 | .92983 | | | 10 | .85590 | 9 | .99591 | | |
| 11 | .99560 | | | 11 | .96615 | | | 11 | .91519 | 10 | .99925 | | |
| | | | | 12 | .98555 | | | 12 | .95306 | 11 | .99989 | | |
| | | | | 13 | .99381 | | | 13 | .97470 | | | | |
| | | | | 14 | .99794 | | | 14 | .98769 | | | | |
| | | | | 15 | .99959 | | | 15 | .99466 | | | | |
| | | | | | | | | 16 | .99789 | | | | |
| | | | | | | | | 17 | .99921 | | | | |
| | | | | | | | | 18 | .99977 | | | | |

| (6, 24, $h$/24) | | (7, 14, $h$/14) | | (7, 28, $h$/28) | | (8, 12, $h$/24) | | (8, 16, $h$/16) | | (8, 32, $h$/32) | | (9, 12, $h$/36) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ |
| 3 | .00123 | 2 | .01881 | 3 | .00033 | 7 | .40465 | 2 | .00892 | 3 | .00009 | 12 | .57284 |
| 4 | .02632 | 3 | .16511 | 4 | .01162 | 8 | .55223 | 3 | .11232 | 4 | .00508 | 13 | .65542 |
| 5 | .10183 | 4 | .43014 | 5 | .05882 | 9 | .67596 | 4 | .34442 | 5 | .03364 | 14 | .70465 |
| 6 | .23167 | 5 | .65870 | 6 | .15755 | 10 | .78579 | 5 | .57501 | 6 | .10610 | 15 | .79433 |
| 7 | .39022 | 6 | .82383 | 7 | .29497 | 11 | .85044 | 6 | .76578 | 7 | .22079 | 16 | .84105 |
| 8 | .54651 | 7 | .91730 | 8 | .44440 | 12 | .90933 | 7 | .87443 | 8 | .35783 | 17 | .88617 |
| 9 | .67067 | 8 | .96687 | 9 | .57474 | 13 | .94434 | 8 | .94207 | 9 | .48770 | 18 | .92151 |
| 10 | .77301 | 9 | .98822 | 10 | .68883 | 14 | .96801 | 9 | .97576 | 10 | .60778 | 19 | .93937 |
| 11 | .85137 | 10 | .99661 | 11 | .78160 | 15 | .98179 | 10 | .99124 | 11 | .71051 | 20 | .95856 |
| 12 | .90678 | 11 | .99917 | 12 | .85191 | 16 | .99138 | 11 | .99720 | 12 | .79265 | 21 | .97396 |
| 13 | .94311 | 12 | .99986 | 13 | .90208 | 17 | .99524 | 12 | .99927 | 13 | .85487 | 22 | .98204 |
| 14 | .96738 | 13 | .99998 | 14 | .93811 | 18 | .99802 | 13 | .99984 | 14 | .90198 | 23 | .98803 |
| 15 | .98242 | | | 15 | .96254 | 19 | .99913 | 14 | .99998 | 15 | .93600 | 24 | .99281 |
| 16 | .99101 | | | 16 | .97818 | 20 | .99968 | 15 | 1.00000 | 16 | .95948 | 25 | .99538 |
| 17 | .99562 | | | 17 | .98775 | 21 | .99986 | 16 | 1.00000 | 17 | .97512 | 26 | .99752 |
| 18 | .99806 | | | 18 | .99351 | | | | | 18 | .98534 | 27 | .99857 |
| 19 | .99923 | | | 19 | .99675 | | | | | 19 | .99170 | 28 | .99901 |
| 20 | .99972 | | | 20 | .99846 | | | | | 20 | .99547 | 29 | .99956 |
| | | | | | | | | | | | | 30 | .99986 |

TABLE 2—*Concluded*

| (9, 15, $h$/45) | | (9, 18, $h$/18) | | (9, 36, $h$/36) | | (10, 15, $h$/30) | | (10, 20, $h$/20) | | (10, 40, $h$/40) | | (12, 15, $h$/60) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ |
| 14 | .51355 | 2 | .00420 | 3 | .00002 | 8 | .40712 | 2 | .00020 | 3 | .00001 | 21 | .72675 |
| 15 | .60637 | 3 | .07584 | 4 | .00220 | 9 | .53773 | 3 | .05090 | 4 | .00095 | 22 | .76664 |
| 16 | .67583 | 4 | .27372 | 5 | .01910 | 10 | .66165 | 4 | .21624 | 5 | .01078 | 23 | .81761 |
| 17 | .73473 | 5 | .49960 | 6 | .07091 | 11 | .74789 | 5 | .43090 | 6 | .04711 | 24 | .84921 |
| 18 | .78605 | 6 | .69439 | 7 | .16403 | 12 | .82837 | 6 | .63135 | 7 | .12214 | 25 | .87374 |
| 19 | .83420 | 7 | .82756 | 8 | .28596 | 13 | .88189 | 7 | .77861 | 8 | .22717 | 26 | .89897 |
| 20 | .86713 | 8 | .91206 | 9 | .41073 | 14 | .92260 | 8 | .87819 | 9 | .34383 | 27 | .92166 |
| 21 | .89937 | 9 | .95874 | 10 | .53221 | 15 | .95017 | 9 | .93771 | 10 | .46325 | 28 | .93902 |
| 22 | .92671 | 10 | .98270 | 11 | .64099 | 16 | .97042 | 10 | .97095 | 11 | .57482 | 29 | .94928 |
| 23 | .94103 | 11 | .99344 | 12 | .73196 | 17 | .98187 | 11 | .98755 | 12 | .67187 | 30 | .96020 |
| 24 | .95820 | 12 | .99783 | 13 | .80416 | 18 | .98997 | 12 | .99520 | 13 | .75196 | 31 | .97033 |
| 25 | .97027 | 13 | .99937 | 14 | .86114 | 19 | .99449 | 13 | .99833 | 14 | .81737 | 32 | .97813 |
| 26 | .97769 | 14 | .99985 | 15 | .90427 | 20 | .99717 | 14 | .99949 | 15 | .86877 | 33 | .98297 |
| 27 | .98515 | 15 | .99997 | 16 | .93569 | 21 | .99853 | 15 | .99986 | 16 | .90780 | 34 | .98692 |
| 28 | .98962 | | | 17 | .95790 | | | 16 | .99997 | 17 | .93668 | 35 | .99045 |
| 29 | .99272 | | | 18 | .97333 | | | | | 18 | .95767 | 36 | .99317 |
| 30 | .99448 | | | 19 | .98362 | | | | | 19 | .97242 | 37 | .99494 |
| 31 | .99637 | | | 20 | .99024 | | | | | 20 | .98249 | 38 | .99650 |
| 32 | .99731 | | | | | | | | | | | 39 | .99754 |
| 33 | .99799 | | | | | | | | | | | 40 | .99824 |
| 34 | .99864 | | | | | | | | | | | 41 | .99870 |

| (12, 16, $h$/48) | | (12, 18, $h$/36) | | (12, 20, $h$/60) | | (15, 20, $h$/60) | | (16, 20, $h$/80) | |
|---|---|---|---|---|---|---|---|---|---|
| $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ | $h$ | $\alpha$ |
| 14 | .53765 | 9 | .41949 | 16 | .49357 | 16 | .56529 | 25 | .74728 |
| 15 | .63564 | 10 | .54635 | 17 | .56236 | 17 | .63688 | 26 | .78477 |
| 16 | .70027 | 11 | .64593 | 18 | .62623 | 18 | .70099 | 27 | .81939 |
| 17 | .76291 | 12 | .73908 | 19 | .68783 | 19 | .74983 | 28 | .84849 |
| 18 | .81604 | 13 | .80850 | 20 | .73647 | 20 | .79723 | 29 | .87017 |
| 19 | .85183 | 14 | .86428 | 21 | .78375 | 21 | .83880 | 30 | .89148 |
| 20 | .88682 | 15 | .90549 | 22 | .82582 | 22 | .87092 | 31 | .91112 |
| 21 | .91630 | 16 | .93737 | 23 | .85490 | 23 | .89808 | 32 | .92759 |
| 22 | .93670 | 17 | .95829 | 24 | .88521 | 24 | .92103 | 33 | .94026 |
| 23 | .95309 | 18 | .97374 | 25 | .90937 | 25 | .93891 | 34 | .95111 |
| 24 | .96637 | 19 | .98371 | 26 | .92760 | 26 | .95393 | 35 | .96069 |
| 25 | .97573 | 20 | .99029 | 27 | .94447 | 27 | .96528 | 36 | .96864 |
| 26 | .98335 | 21 | .99429 | 28 | .95708 | 28 | .97389 | 37 | .97489 |
| 27 | .98848 | | | 29 | .96699 | 29 | .98044 | 38 | .98032 |
| 28 | .99190 | | | 30 | .97431 | 30 | .98586 | 39 | .98458 |
| 29 | .99423 | | | 31 | .98093 | 31 | .99007 | 40 | .98790 |
| 30 | .99624 | | | 32 | .98554 | | | 41 | .99051 |
| 31 | .99775 | | | 33 | .98915 | | | | |
| | | | | 34 | .99216 | | | | |

# THE STOCHASTIC INDEPENDENCE OF SYMMETRIC AND HOMOGENEOUS LINEAR AND QUADRATIC STATISTICS

By Eugene Lukacs

*National Bureau of Standards*

**1. Summary.** The following theorem is proved. *If a univariate distribution has moments of first and second order and admits a homogeneous and symmetric quadratic statistic $Q$ which is independently distributed of the mean of a sample of $n$ drawn from this distribution, then it is either the normal distribution* ($Q$ is then proportional to the variance) *or the degenerate distribution* (in this case no restriction is imposed on $Q$) *or a step function with two symmetrically located steps* (in this case $Q$ is the sum of the squared observations). *The converse of this statement is also true.*

**2. Introduction.** It is known that the distributions of the mean and of the variance in samples from a continuous population are stochastically independent if and only if the parent distribution is normal. This theorem was proven independently by several authors [1], [2], [3].

The question arises whether there are any distributions having the property that the sampling distributions of the mean and of a symmetric and homogeneous quadratic statistic are stochastically independent. This is the problem to be discussed in the present paper; all distributions with this property will be determined and also the corresponding quadratic statistics. In this discussion we consider a constant as stochastically independent of any random variable.

In an earlier paper [3], dealing with the independence of the mean and the variance, the existence of a frequency function was assumed. To obtain all possible distribution functions it is necessary to refrain from this assumption and to express the formulae in terms of the cumulative distribution functions. Otherwise the derivation of the differential equation of the characteristic function (Section 3) resembles the reasoning of the preceding paper. In Section 4 and Section 5 this differential equation is discussed and the various possible solutions are determined.

**3. The differential equation for the characteristic function.** We consider a univariate population with a cumulative distribution function $F(x)$. Let $x_1, x_2, \cdots, x_n$ be $n$ independent observations of the variate $x$. The cumulative distribution function of the sample is then given by $\Phi(x_1, x_2, \cdots, x_n) = F(x_1) F(x_2) \cdots F(x_n)$. Let us set $L = \sum_{j=1}^{n} x_j$ and $S = \sum_{j=1}^{n} x_j^2$. Any symmetric and homogeneous quadratic statistic $Q$ can then be expressed in terms of $L$ and $S$:

$$Q = aL^2 + bS. \tag{1}$$

The problem under investigation is the determination of all cumulative distribution functions $F(x)$ and statistics $Q$ for which the sampling distributions of $Q$ and of the mean $\bar{x} = \sum_{j=1}^{n} x_j/n$ (or equivalently of $L = n\bar{x}$) are stochastically independent.

The characteristic function of the distribution $F(x)$ is given by

$$\psi(u) = \int_{-\infty}^{+\infty} e^{iux}\, dF(x). \tag{2}$$

The characteristic function of the statistic $L$ is

$$\phi_1(t) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} e^{itL}\, d\Phi(x_1 \cdots x_n) = \prod_{j=1}^{n} \int_{-\infty}^{+\infty} e^{itx_j}\, dF(x_j) = [\psi(t)]^n. \tag{3.1}$$

Similarily the characteristic function of the statistic $Q$ is given by

$$\phi_2(v) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} e^{ivQ}\, d\Phi(x_1, \cdots, x_n), \tag{3.2}$$

and the characteristic function of the joint distribution of $L$ and $Q$ by

$$\phi(t, v) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} e^{itL+ivQ}\, d\Phi(x_1, \cdots, x_n). \tag{3.3}$$

The statistics $Q$ and $L$ are independently distributed if and only if

$$\phi(t, v) = \phi_1(t)\phi_2(v). \tag{4}$$

Differentiating (4) with respect to $v$ and then putting $v = 0$ one has

$$\left.\frac{\partial \phi}{\partial v}\right|_{v=0} = \phi_1(t) \left.\frac{\partial \phi_2}{\partial v}\right|_{v=0} = [\psi(t)]^n \left.\frac{\partial \phi_2}{\partial v}\right|_{v=0},$$

or from (3.1), (3.2) and (3.3)

$$\begin{aligned}\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} Qe^{itL}\, d\Phi &= \left\{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} e^{itL}\, d\Phi\right\}\left\{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} Q\, d\Phi\right\} \\ &= [\psi(t)]^n \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} Q\, d\Phi.\end{aligned} \tag{5}$$

Any distribution function $F(x)$ and any statistic $Q$ which satisfies (4) also satisfies condition (5). It is worthwhile to remark that (5) implies certain restrictions on the statistic $Q$.

In the following we determine the distributions $F(x)$ and statistics $Q$ which satisfy (5). This is done by first transforming (5) into a differential equation for the characteristic function $\psi(t)$ and by finding also a statistic $Q$ using (5). Finally we shall investigate which of these distributions and statistics will also satisfy condition (4) and constitute hereby a solution to our problem.

We denote the left-hand side of (5) by $I(t)$. Substituting for $Q$ from (1) we obtain after a simple computation

$$I(t) = -n[\psi(t)]^{n-2}\left[(a+b)\psi(t)\frac{d^2\psi}{dt^2} + a(n-1)\left(\frac{d\psi}{dt}\right)^2\right].$$

This permits one to reduce (5) to a differential equation for the characteristic function

$$(a+b)\left[\frac{1}{\psi(t)}\frac{d^2\psi}{dt^2}\right] + a(n-1)\left[\frac{1}{\psi(t)}\frac{d\psi}{dt}\right]^2 = A, \tag{6}$$

where

$$A = -[(a+b)\alpha_2 + (n-1)a\alpha_1^2], \tag{6.1}$$

with the initial conditions

$$\psi(0) = 1; \qquad \psi'(0) = i\alpha_1. \tag{6.2}$$

Here

$$\alpha_r = \int_{-\infty}^{+\infty} x^r\, dF(x).$$

From (6) and (6.2) we obtain easily first an equation for the cumulant generating function $g(t) = \ln \psi(t)$ and then the following differential equation for $h(t) = \frac{dg}{dt}$:

$$(a+b)\frac{dh}{dt} + (na+b)[h(t)]^2 = A, \tag{7}$$

with

$$h(0) = i\alpha_1. \tag{7.1}$$

**4. Discussion of the differential equation (7) if $A \neq 0$.** We have to distinguish three possibilities:

$$A \neq 0, \qquad na + b = 0, \qquad a + b \neq 0; \tag{8.1}$$

$$A \neq 0, \qquad na + b \neq 0, \qquad a + b \neq 0; \tag{8.2}$$

$$A \neq 0, \qquad na + b \neq 0, \qquad a + b = 0. \tag{8.3}$$

We first discuss case (8.1); equation (7) reduces to $(a+b)dh/dt = A$; from (6.1) and $na + b = 0$ it follows that

$$\frac{dh}{dt} = -\sigma^2 = -(\alpha_2 - \alpha_1^2).$$

Integrating this equation and considering that $h(t) = dg/dt$ we obtain using the initial condition (7.1), $g(t) = -\frac{1}{2}\sigma^2 t^2 + i\alpha_1 t$. This is the cumulant generating

function of the normal distribution. Since $b = -na$ we find $Q = -an^2s^2$. Here $s^2 = \sum_{j=1}^{n} x_j^2/n - (\sum_{j=1}^{n} x_j)^2/n^2$. In case (8.1) the parent distribution is normal and the statistic $Q$ is—except for a constant factor—the sample variance. Conversely it is known that for the normal distribution variance and mean are stochastically independent.

We start the discussion of case (8.2) with the almost trivial case when $h(t)$ reduces to a constant, and delay dealing with the general case. If $h(t)$ is a constant and if (8.2) holds then equation (7) is satisfied if $h^2(t) = A/(na + b) = \text{const.}$ Then $h(t) = h(0) = i\alpha_1$, and since $dg/dt = h(t)$, $g(0) = 0$ we obtain $g(t) = i\alpha_1 t$ and $\psi(t) = e^{i\alpha_1 t}$. The function $\psi(t) = e^{i\alpha_1 t}$ is the characteristic function of the degenerate distribution

$$\varepsilon(x - \alpha_1) = \begin{cases} 1 & \text{if} \quad x \geqq \alpha_1, \\ 0 & \text{if} \quad x < \alpha_1. \end{cases}$$

It is easy to show that for this distribution any symmetric and homogeneous quadratic statistic is independent of the mean.

We proceed with the discussion of case (8.2) by assuming that

$$h^2(t) \not\equiv \frac{A}{na + b}. \tag{9}$$

Equation (7) may then be written as

$$(a + b)\frac{dh}{dt} = A\left[1 - \frac{na + b}{A}h^2(t)\right].$$

We integrate this equation and then compute the function $\psi(t)$ from the relation $h(t) = dg/dt = d/dt \ln \psi(t)$, with $\psi(0) = 1$.
Thus we obtain

$$\psi(t) = [pe^{\frac{1}{2}\beta t} + qe^{-\frac{1}{2}\beta t}]^{\lambda}, \tag{10}$$

with

$$\lambda = \frac{a + b}{na + b}, \qquad p = \frac{C}{C + 1}, \qquad q = \frac{1}{C + 1},$$

$$\beta = \frac{2}{\lambda}\sqrt{\frac{A}{na + b}} = \frac{2}{\lambda}\sqrt{(\lambda - 1)\alpha_1^2 - \lambda\alpha_2}, \tag{10.1}$$

$$C = \frac{\lambda\beta + 2i\alpha_1}{\lambda\beta - 2i\alpha_1}.$$

In these formulae $a$ and $b$ and therefore also $A$, $\beta$, $C$ as well as $\lambda$, $p$ and $q$ may be functions of $n$.

In the following we have to distinguish two cases. We first assume that

$$\lambda\alpha_2 > (\lambda - 1)\alpha_1^2. \tag{8.2.1}$$

Then $\beta$ is purely imaginary so that $C$, $p$ and $q$ are real numbers. Since $p + q = 1$, at least one of the two numbers $p$ and $q$ must be positive. If both $p$ and $q$

are positive then it can be shown by computing the inverse Fourier transform of (10) that (10) can be a characteristic function only if $\lambda$ is a positive integer. The corresponding distribution is in this case a binomial distribution. If one of the numbers $p$ and $q$ is positive while the other is negative then we see easily that $\lambda < 0$. In this case (10) is the characteristic function of a negative binomial distribution.

The binomial as well as the negative binomial distribution have the property that the random variable assumes an extreme value with a positive probability. Let $L_0$ be the extreme value which the statistic $L$ may assume; we have then $P(L = L_0) > 0$. If $L = L_0$ then the value of $Q$ is completely determined. Let us denote it by $Q_0$. Therefore the conditional probability $P(Q = Q_0 \mid L = L_0) = 1$; on account of the assumed stochastic independence of $Q$ and $L$ this means that $P(Q = Q_0) = 1$ so that $Q$ has necessarily a degenerate distribution. We see then from (1) that $x^2$ has also a degenerate distribution. We conclude that under the conditions (8.2) and (8.2.1) the case $p > 0$, $q < 0$ (or $p < 0$, $q > 0$) can not occur. The distribution of $x$ has therefore necessarily the form

$$F(x) = p\varepsilon(x - \xi) + q\varepsilon(x + \xi), \tag{10.2}$$

where

$$p > 0, \qquad q > 0, \qquad p + q = 1. \tag{10.3}$$

If we take $a = 0$, then $Q = bS$. From (10.2) it is seen that $S = n\alpha_2 = n\xi^2$ with probability one and hence $Q = bS$ is independent of $L$. Furthermore, if $n > 1$, $aL^2 + bS$ is not independent of $L$ unless $p = 0$ or $p = 1$, and hence $a$ must be zero and $Q = b\sum_{j=1}^{n} x_j^2$. On the other hand it is easily seen that for the distribution (10.2) the statistics $Q = \sum_{j=1}^{n} x_j^2$ and $L = \sum_{j=1}^{n} x_j$ are stochastically independent. This completes the discussion of the case where (8.2) and (8.2.1) hold.

We still have to consider the possibility that (8.2) is valid while

$$\lambda\alpha_2 \leqq (\lambda - 1)\alpha_1^2. \tag{8.2.2}$$

In this case $\beta$ is real. Clearly, we must have $\lambda < 0$ since otherwise $\psi(t)$ would not be bounded. We put $\lambda = -\mu(\mu > 0)$ and write

$$\psi(t) = [pe^{i\beta t} + qe^{-i\beta t}]^{-\mu}. \tag{11}$$

Here $p$ and $q$ must be both positive since otherwise $\psi(t)$ would have a real pole. It is also seen easily that $\psi(t)$ has a maximum for a value $t_0$ given by $e^{\beta t_0} = q/p$. Furthermore, if $p \neq q$ we see that $\psi(t_0) = [2\sqrt{pq}]^{-\mu} > 1$, while $\psi(t_0) = 1$ if $p = q = \frac{1}{2}$. If $\psi(t)$ is a characteristic function we must therefore have $p = q = \frac{1}{2}$ so that (11) reduces to

$$\psi(t) = \left[\frac{1}{\cosh\dfrac{\beta t}{2}}\right]^{\mu} = \left[\prod_{j=1}^{\infty} \frac{1}{1 + \dfrac{\beta^2 t^2}{(2j-1)^2\pi^2}}\right]^{\mu}. \tag{11.1}$$

The representation of $\psi(t)$ as an infinite product shows that (11.1) is a characteristic function. We show finally that for the distributions determined by (11.1) no homogeneous and symmetric quadratic statistic $Q$ exists which is independent of the statistic $L$. This is proven by demonstrating that $E(Q^2L^2) - E(Q^2)E(L^2)$ can not be zero for any distribution with the characteristic function (11.1). The symbol $E$ denotes here as usual the expected value.

By a somewhat tedious but elementary computation one obtains from (11.1) and (1)

$$E(Q^2L^2) - E(Q^2)\,E(L^2) = 4n\mu\left(\frac{\beta}{2}\right)^6 a^2\Big\{(\mu+1)[(n+2)\mu+4]\left(\frac{b}{a}\right)^2 + [(n^2+5n)\mu^2+(6n+8)\mu+8]\left(\frac{b}{a}\right) + (n\mu+1)(3n\mu+4)\Big\}.$$

From $\mu = -\lambda$ and (10.1) we see that $b/a = -(n\mu+1)/(\mu+1)$. If we substitute this into the previous equation and simplify we obtain finally

$$E(Q^2L^2) - E(Q^2)E(L^2) = \frac{n\mu^2a^2(n\mu+1)(n-1)\beta^6}{8(\mu+1)}. \tag{11.2}$$

Since $\mu > 0$, this can be zero only if $n = 1$. Then either $Q \equiv 0$ or $a + b \neq 0$. By (10.1) $n = 1$ and $a + b \neq 0$ imply $\lambda = 1$, which contradicts $\mu = -\lambda > 0$. The distribution functions (11.1), derived under the assumptions (8.2) and (8.2.2), do not therefore yield a solution of our problem.

We next consider the case (8.3) by assuming $A \neq 0$, $a + b = 0$, $na + b \neq 0$. Equation (7) then reduces to $[h(t)]^2 = -\alpha_1^2$ which leads to $\psi(t) = e^{\pm i\alpha_1 t}$ and $Q = 2a\sum_{j=1}^{n-1}\sum_{k=j+1}^{n} x_jx_k$. We see therefore that case (8.3) leads to no new distribution as the degenerate distribution appears already in the preceding discussion.

**5. Discussion of the differential equation if $A = 0$.** In this paragraph we assume

$$A \equiv (a+b)\alpha_2 + (n-1)a\alpha_1^2 = 0, \tag{12.1}$$

and rewrite equation (7) as

$$(a+b)\frac{dh}{dt} + (na+b)\,h^2(t) = 0. \tag{12.2}$$

We start the discussion of equations (12.1) and (12.2) with four cases which lead either to an already known solution for the characteristic function or to a trivial solution for the statistic $Q$. We assume first

$$a = 0. \tag{12.1.1}$$

From (12.1.1) and (12.1) we see $b\alpha_2 = 0$. If $b = 0$ we obtain the improper statistic $Q \equiv 0$, which is independent of any other statistic whatever be the parent

distribution. If $\alpha_2 = 0$, we obtain the characteristic function $\psi(t) = 1$. The same solutions (either $\psi(t) = 1$ or $Q \equiv 0$) are also derived from (12.1), (12.2) and

$$a + b = 0, \tag{12.1.2}$$

or from (12.1), (12.2) and

$$\alpha_1 = 0. \tag{12.1.3}$$

If we assume

$$na + b = 0, \tag{12.1.4}$$

we obtain in a similar manner either the statistic $Q \equiv 0$ or the characteristic function $\psi(t) = e^{it\alpha_1}$.

We see therefore that if (12.1) holds the cases $a = 0$, $a + b = 0$, $\alpha_1 = 0$, $na + b = 0$ lead again either to a degenerate distribution which was fully treated in the preceding section or to the improper statistic $Q \equiv 0$, which is a trivial solution of our problem.

In the following we therefore discuss equation (12.2) only under the assumption that the four inequalities

$$a \neq 0, \qquad na + b \neq 0, \qquad a + b \neq 0, \qquad \alpha_1 \neq 0 \tag{12.3}$$

hold. In this case (12.2) may be written as $\frac{d}{dt}\frac{1}{h(t)} = -\frac{\sigma^2}{\alpha_1^2}$, where $\sigma^2 = \alpha_2 - \alpha_1^2$ is the variance of the parent distribution. If we integrate this with due regard to the initial conditions we obtain

$$\psi(t) = \left[1 - \frac{\sigma^2 it}{\alpha_1}\right] - \frac{\alpha_1^2}{\sigma^2}. \tag{13}$$

This is the characteristic function of a gamma-distribution with parameters $\alpha = \alpha_1/\sigma^2$, $\lambda = \alpha_1^2/\sigma^2$. Its frequency function is given by

$$f(x, \alpha, \lambda) = \begin{cases} \dfrac{\alpha^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-\alpha x} & \text{for } x > 0, \\ 0 & \text{for } x \leqq 0, \end{cases} \qquad \text{if } \alpha_1 > 0 \tag{13.1}$$

and in case $\alpha_1 < 0$ by

$$g(x, \alpha, \lambda) = \begin{cases} 0 & \text{for } x \geqq 0, \\ \dfrac{(-\alpha)^\lambda}{\Gamma(\lambda)} (-x)^{\lambda-1} e^{-\alpha x} & \text{for } x < 0. \end{cases} \tag{13.2}$$

Equation (12.1), which led to the characteristic function (13) of the gamma-distribution, determines also the statistic $Q$. It is easily seen from (12.1) that $\sigma^2/\alpha_1^2 = -(na + b)/(a + b) = 1/\lambda$, therefore

$$\frac{b}{a} = -\frac{n\lambda + 1}{\lambda + 1}. \tag{14}$$

We show finally that the gamma distribution determined by (13) as a solution of the differential equation (12.2) does not lead to a solution of our problem since no statistic $Q$ exists which for any sample size is independently distributed of $L$. Suppose $n > 1$. Then $|Q| = |a(L^2 - (n\lambda + 1)/(\lambda + 1)S)| \leqq |a|(L^2 + (n\lambda + 1)/(\lambda + 1)S)$. But $S \leqq L^2$ with probability one since all $x$'s are of one sign with probability one. Hence $|Q| \leqq |a|[((n + 1)\lambda + 2)/(\lambda + 1)]L^2$ with probability one. The range of the values of $Q$ is unbounded, hence $Q$ can not be independent of $L$.

## REFERENCES

[1] R. C. Geary, "Distribution of Student's ratio for non-normal samples," *Jour. Roy. Stat. Soc. Suppl.*, Vol. 3 (1936), pp. 178–184.

[2] T. Kawata and H. Sakamoto, "On the characterization of the normal population by the independence of the sample mean and the sample variance," *Jour. Math. Soc. Japan*, Vol. 1 (1949), pp. 111–115.

[3] E. Lukacs, "A characterization of the normal distribution," *Annals of Math. Stat.*, Vol. 13 (1942), pp. 91–93.

# ON A TEST FOR HOMOGENEITY AND EXTREME VALUES

By D. A. Darling

*University of Michigan*

**Summary.** Let $x_1, x_2, \cdots, x_n$ be positive, identically distributed, independent random variables. It is of some statistical interest to study the distribution of $z_n = (x_1 + x_2 + \cdots + x_n)/\max(x_1, x_2, \cdots, x_n)$. In this paper we give its characteristic function and in a few cases its distribution. A limiting distribution of fairly wide applicability is given in the last section.

**1. Introduction.** Suppose $x_1, x_2, \cdots, x_n$ are positive, independent observations coming from populations with cumulative distribution functions $F(\sigma_1 x)$, $F(\sigma_2 x), \cdots, F(\sigma_n x)$ respectively, where $\sigma_i > 0$ and $F(x)$ is some known continuous cdf with $F(0) = 0$. It is of some practical importance to devise a test for the hypothesis $H$ that $\sigma_1 = \sigma_2 = \cdots = \sigma_n$. For instance, in the analysis of variance it is desirable to know if a set of $n$ independent sample variances are "homogeneous," or if to the contrary they come from normal populations with different variances.

Let $t_n = x_1 + x_2 + \cdots + x_n$ and $m_n = \max(x_1, x_2, \cdots, x_n)$, and put $u_n = m_n/t_n$. A possible test consists of rejecting $H$ when $u_n$ exceeds a certain value. In 1929 R. A. Fisher [1] used this test in determining whether the largest amplitude in a harmonic analysis was "significantly" large, $H$ being the hypothesis that each $x_i$ was independently distributed as $\sigma^2\chi^2$ with 2 degrees of freedom. In 1941 Cochran [2] considered an extension to $k$ degrees of freedom. Other writers have taken the ratio of an extreme variance to an independent estimate of the variance as a test for homogeneity and outlying observations (cf. Nair [3] and Finney [4]). This test is also used implicitly in quality control work to test the equality of scale parameters; here $x_1, x_2, \cdots, x_n$ are each independently distributed as the range of a sample of size $k$.

A test of the nature described above is supposed to have good power against alternatives which consist of the occurrence of just one anomalously large value of $\sigma$, whereas for other alternatives (in the case of testing variances) the well known test of Bartlett (which is essentially the likelihood ratio test) is perhaps superior. This test might also be employed to determine the equality of location parameters in a set of cdf's, and generally to test the significance of extreme values. Since, under $H$, the distribution of $u_n$ is independent of the common population variances (or is "Studentized" in the terminology of Hartley [5]), it is superior, for many purposes, to tests based on other order statistics (cf. Pearson and Chandrasekhar [6]).

The distribution of $u_n$ when $H$ is true is not, in general, known. In this note we consider the distribution of $z_n = 1/u_n = t_n/m_n$. Of course, the test of $H$ using $z_n$ is identical to the test using $u_n$, and if the critical rejection region using $u_n$ is $u_n > \alpha$, then the critical region using $z_n$ is $z_n < 1/\alpha$. We give the

characteristic function of $z_n$ and show how, in some cases, it can be used to determine the rejection region.

**2. The distribution of $Z_n$.** Let $x_1, x_2, \cdots, x_n, t_n$ and $m_n$ be the variables described in the preceding section, and let $z_n = t_n/m_n$. Suppose the variables $x_i$ have the common density $\phi(x)$, $x > 0$. Then for the characteristic function of $z_n$ we have the following formula, proven in [7]:

$$\xi_n(t) = E(e^{itz_n}) = ne^{it} \int_0^\infty \left(\beta \int_0^1 e^{it\alpha} \phi(\alpha\beta)\, d\alpha\right)^{n-1} \phi(\beta)\, d\beta. \tag{2.1}$$

In certain cases it is possible to invert this Fourier transform, and so obtain the distribution of $z_n$, and in any case we can find all moments of $z_n$ in terms of quadratures. For example, the mean is

$$\mu = E(z_n) = \frac{1}{i}\xi_n'(0) = 1 + n(n-1) \int_0^\infty (F(\beta))^{n-2} \left\{\int_0^1 \phi(\beta u)\, du - \phi(\beta)\right\} \phi(\beta)\, d\beta,$$

where $F(\beta)$ is the cdf, that is, $F(\beta) = \int_0^\beta \phi(x)\, dx$.

We consider a few examples.

(a) Suppose the variables $x_i$ are uniformly distributed over (0, 1)—or for that matter uniformly distributed over any finite interval $(0, a)$, since the distribution of $z_n$ is independent of $a$. Then (2.1) becomes

$$\xi_n(t) = e^{it}\left(\frac{e^{it} - 1}{it}\right)^{n-1}$$

and we have the rather surprising result the distribution of $z_n$ is the same as the distribution of $1 + y_1 + y_2 + \cdots + y_{n-1}$, where $y_1, y_2, \cdots y_{n-1}$ are independent and uniformly distributed over (0, 1). This fact may also be readily deduced from certain properties of order statistics of uniformly distributed variables which have been studied recently by Malmquist [8].

(b) Let the variables have the density

$$\phi(x) = \frac{1}{\Gamma\left(\frac{k}{2}\right)} e^{-x} x^{\frac{k}{2}-1}, \tag{2.2}$$

which is the density of $2\chi^2$ with $k$ degrees of freedom. For variables which have the density $\sigma^2\chi^2$ the characteristic function of $z_n$ is given by (2.1) with $\phi(x)$ as in (2.2). If $k$ is an even integer we can find the distribution function of $z_n$ explicitly. Let $k = 2r + 2$ where $r$ is an integer $\geq 0$. Then (2.2) becomes

$$\phi(x) = \frac{1}{r!} e^{-x} x^r,$$

and if we work with the Laplace transform $\xi_n(it)$ rather than $\xi_n(t)$, (2.1) becomes

$$\xi_n(it) = E(e^{-itz_n}) = ne^{-t}\int_0^\infty \psi^{n-1}(\beta)\,\frac{1}{r!}\,e^{-\beta}\beta^r\,d\beta, \tag{2.3}$$

where

$$\psi(\beta) = \frac{\beta}{r!}\int_0^1 (\alpha\beta)^r e^{-(\beta+t)\alpha}\,d\alpha,$$

$$= \left(\frac{\beta}{\beta+t}\right)^{r+1}(1 - e^{-(\beta+t)}P_r(\beta+t)),$$

in which $P_r(x)$ is the polynomial $P_r(x) = \sum_{j=0}^r x^j/j!$. Now $\xi_n(it)/t$ is the Laplace transform of the cumulative distribution function of $z_n$, and (2.3) can be rewritten

$$\frac{\xi_n(it)}{t} = \frac{n}{r!}\int_1^\infty \left(1 - \frac{1}{u}\right)^{(n-1)(r+1)}(u-1)^r\,(1 - e^{-ut}P_r(ut))^{n-1}t^r e^{-ut}\,du.$$

Letting

$$f(u) = \begin{cases} \left(1 - \dfrac{1}{u}\right)^{(n-1)(r+1)}(u-1)^r, & u > 1, \\ 0, & u \leqq 1, \end{cases}$$

and $f^{(k)}(u) = 0$ for all $k \geqq 0$ and $u \leqq 1$ we have

$$\frac{\xi_n(it)}{t} = \frac{1}{r!}\sum_{j=1}^n \binom{n}{j}(-1)^{j+1}\int_j^\infty f\left(\frac{u}{j}\right)\left[P_r\left(\frac{ut}{j}\right)\right]^{j-1}e^{-ut}t^r\,du,$$

and putting $(P_r(x))^i = \sum_{\nu=0}^{ir} a_\nu^{(i)}x^\nu$,

$$\frac{\xi_n(it)}{t} = \frac{1}{r!}\sum_{j=1}^n \binom{n}{j}(-1)^{j-1}\int_0^\infty f\left(\frac{u}{j}\right)\left\{\sum_{\nu=0}^{(j-1)r} a_\nu^{(j-1)}\left(\frac{u}{j}\right)^\nu t^{r+\nu}\right\}e^{-ut}\,du,$$

and this expression can be now directly inverted termwise to give

$$Pr\{z_n \leqq x\} = \frac{1}{r!}\sum_{j=1}^n \binom{n}{j}(-1)^{j-1}\left\{\sum_{\nu=0}^{r(j-1)} a_\nu^{(j-1)}\,\frac{d^{r+\nu}}{dx^{r+\nu}}f\left(\frac{x}{j}\right)\left(\frac{x}{j}\right)^r\right\},$$

or finally

$$Pr\{z_n \leqq x\} = \frac{1}{r!}\sum_{1\leqq j<x} \binom{n}{j}(-1)^{j-1}\sum_{\nu=0}^{r(j-1)} a_\nu^{(j-1)} \cdot \frac{d^{r+\nu}}{dx^{r+\nu}}\left(1 - \frac{j}{x}\right)^{(n-1)(r+1)}\left(\frac{x}{j} - 1\right)^r\left(\frac{x}{j}\right)^r, \qquad 1 \leqq x \leqq n. \tag{2.4}$$

If we put $r = 0$, and calculate the distribution of $u_n = 1/z_n$, (2.4) becomes

$$Pr\{u_n > x\} = \sum_{1\leqq j<(1/x)} \binom{n}{j}(-1)^{j-1}(1 - jx)^{n-1}, \quad \frac{1}{n} \leqq x \leqq 1,$$

which is the result of Fisher [1]. Cochran [2] showed how to express the terms of $u_n$, corresponding to (2.4), for an even number of degrees of freedom as multiple integrals of Beta functions, and on the basis of certain approximations gave a table of approximate percentage points for $u_n$. This table, as well as a subsequent table by Eisenhart, Hastay and Wallis [9], appears to be subject to certain indeterminate errors (cf. [9] ch. 15).

**3. A limiting distribution.** In example (b) above it is possible to find a limiting distribution for $z_n$ when $n \to \infty$ which may be of use when $n$ is large. (The limiting distribution for $u_n = 1/z_n$ given by Cochran [2] appears to be erroneous owing to an oversight concerning the dependence of a set of random variables.) We remove the restriction that $r$ be a nonnegative integer and merely require $r > -1$. Then

$$\phi(\beta) = \frac{1}{\Gamma(r+1)} e^{-\beta}\beta^r$$

and $\xi_n(t)$ is given by (2.1). Now define $f(\beta) = 1 - F(\beta) = \int_\beta^\infty \phi(x)\,dx$ and

$$\begin{aligned}\psi(\beta) &= 1 - f(\beta) + \beta\int_0^1 (e^{it\alpha} - 1)\phi(\alpha\beta)\,d\alpha \\ &= 1 - f(\beta) + g(\beta),\end{aligned}$$

so that $\xi_n(t) = -ne^{it}\int_0^\infty \psi^{n-1}(\beta)\,df(\beta)$. It is necessary to study the behavior of $g(\beta)$ for large $\beta$, and it is simple to show that

$$\begin{aligned}g(\beta) &= \beta\int_0^1 (e^{it\alpha} - 1)\phi(\alpha\beta)\,d\alpha \\ &= \frac{it(r+1)}{\beta} + \frac{(it)^2}{\beta^2}\frac{(r+1)(r+2)}{2} + o\left(\frac{1}{\beta^2}\right), \qquad \beta \to \infty\end{aligned}$$

for bounded $|t|$.

We next need to get an asymptotic solution to the equation $nf(\beta) = v$ for $\beta$ when $n \to \infty$; that is, we need to solve

$$v = \frac{n}{\Gamma(r+1)}\int_\beta^\infty e^{-x}x^r\,dx$$

for $\beta$ when $n \to \infty$. An asymptotic development similar to that above shows that we have

$$v = \frac{n}{\Gamma(r+1)} e^{-\beta}\beta^r\left(1 + \frac{r}{\beta} + o\left(\frac{1}{\beta}\right)\right)$$

as an equation to solve for $\beta$. After some calculation we find

$$\beta = \log n + r\log\log n + r^2\frac{\log\log n}{\log n} - \log(v\Gamma(r+1)) + O\left(\frac{1}{\log n}\right),$$

and hence

$$\frac{1}{\beta} = \frac{1}{\log n} - r\frac{\log\log n}{\log^2 n} + \frac{\log(v\Gamma(r+1))}{\log^2 n} + o\left(\frac{1}{\log^2 n}\right),$$

so that

$$\psi(\beta) = 1 - f(\beta) + g(\beta) = 1 - \frac{v}{n} + it(r+1)$$
$$\left\{\frac{1}{\log n} - \frac{r\log\log n}{\log^2 n} + \frac{\log(v\Gamma(r+1))}{\log^2 n} + o\left(\frac{1}{\log^2 n}\right)\right\}.$$

Now define numbers $c_n$ and $d_n$ as follows:

$$c_n = (r+1)\frac{n}{\log^2 n},$$

$$d_n = \log n - r\log\log n + \log\Gamma(r+1),$$

and it is simple to verify that

$$\xi_n(t/c_n)e^{-it d_n} = e^{it/c_n}\int_0^n \left(1 - \frac{v}{n} + \frac{it\log v}{n} + o\left(\frac{1}{n}\right)\right)^{n-1} dv,$$

and that when $n \to \infty$ the limit may be taken under the integration sign to give

$$\xi_n(t/c_n)e^{-it d_n} \to \int_0^\infty e^{-v+it\log v}\,dv = \Gamma(it+1).$$

But the expression on the right is simply the characteristic function for the random variable $\log w$, where $w$ has the exponential density $e^{-x}$, and hence

$$\lim_{n\to\infty} Pr\{z_n/c_n - d_n < x\} = Pr\{\log w < x\} = 1 - e^{-e^x}, \qquad -\infty < x < \infty,$$

by the continuity theorem for characteristic functions. It does not follow, of course, that the constants $c_n$ and $d_n$ are the "best" in the sense that they give the "closest" approximation to the limiting distribution function when $n$ is finite.

Finally, then

$$\lim_{n\to\infty} Pr\left\{z_n < (r+1)\frac{n}{\log n} - r(r+1)\frac{n\log\log n}{\log^2 n} + (r+1)\log\Gamma(r+1)\right.$$
$$\left.\cdot\frac{n}{\log^2 n} + (r+1)\frac{n}{\log^2 n}x\right\} = 1 - e^{-e^x}, \qquad -\infty < x < \infty.$$

**4. A general limiting expression.** Following the analysis of Section 3, it is possible to get a general limiting distribution for $z_n$ for quite a broad class of distributions $\phi(x)$. There are two essentially distinct cases.

(a) If $x_i$ is bounded (i.e., if $\phi(x) = 0$ for $x > x_0$) then $m_n = \max(x_1, x_2, \cdots, x_n)$

tends with probability 1 to some sure number $M$ and hence $z_n$ has the same asymptotic distribution as $1/M\ (x_1 + x_2 + \cdots + x_n)$—that is, the Gaussian distribution.

(b) If $\mu = E(x_i)$ exists and $x_i$ is unbounded from above we define, as before $\psi(\beta) = 1 - f(\beta) + g(\beta)$, where

$$g(\beta) = \beta\int_0^1 (e^{i\beta\alpha} - 1)\phi(\alpha\beta)\, d\alpha,$$

so that (2.1) becomes $\xi_n(t) = -ne^{it}\int_0^\infty (\psi(\beta))^{n-1}\, df(\beta)$, $f(\beta)$ being $1 - F(\beta) = \int_\beta^\infty \phi(x)\, dx$ as in Section 3. An asymptotic development for $g(\beta)$ is

$$g(\beta) = \frac{it\mu}{\beta} + o\left(\frac{1}{\beta}\right), \qquad \beta \to \infty,$$

and as before we must solve $nf(\beta) = v$ for $1/\beta$, asymptotically for $n \to \infty$.

Let us suppose that there are constants $a_n$ and $b_n$ such that if $nf(\beta) = v$, $v$ bounded,

(1) $\dfrac{1}{\beta} = a_n + h(v)b_n + f_n(v), \qquad n \to \infty,$

(2) $nb_n \to \infty, \qquad \dfrac{a_n}{nb_n} \to 0, \qquad n \to \infty,$

(3) $f_n(v)/b_n \to 0$ uniformly for $v$ in any closed interval $0 < \epsilon \leqq v \leqq M < \infty$.

Then $h(v)$ is a monotone increasing function, and we consider the normalized variable $(z_n - \mu n a_n)/\mu n b_n$. For its characteristic function we have

$$\xi_n(t/nb_n)e^{-it(a_n/b_n)} = \int_0^n \left(1 - \frac{v}{n} + \frac{it\, h(v)}{n} + o\left(\frac{1}{n}\right)\right)^{n-1} dv \to \int_0^\infty e^{-v+it\,h(v)}\, dv.$$

As a consequence the normalized variable has the limiting distribution of $h(w)$, where $w$ is distributed with an exponential density $e^{-x}$. Hence

$$Pr\left\{\frac{z_n}{nb_n} - \frac{a_n}{b_n} < x\right\} \to Pr\{h(w) < x\} = 1 - e^{-h^{-1}(x)},$$

or

$$Pr\{z_n < \mu n a_n + \mu n b_n x\} \to 1 - e^{-h^{-1}(x)}.$$

## REFERENCES

[1] R. A. Fisher, "Tests of significance in harmonic analysis," *Proc. Roy. Soc. Edinburgh Sect. A*, Vol. 125 (1929), pp. 54–59.

[2] W. G. Cochran, "The distribution of the largest of a set of estimated variances as a fraction of their total," *Annals of Eugenics*, Vol. 11 (1941), pp. 47–52.

[3] K. R. NAIR, "The studentized form of the extreme mean square test in the analysis of variance," *Biometrika*, Vol. 35 (1948), pp. 16–31.

[4] D. J. FINNEY, "The joint distribution of variance ratios based on common error mean square," *Annals of Eugenics*, Vol. 11 (1941), pp. 136–140.

[5] H. O. HARTLEY, "Studentization and large sample theory," *Jour. Roy. Stat. Soc. Suppl.*, Vol. 5 (1938), pp. 80–90.

[6] E. S. PEARSON AND C. CHANDRASEKAR, "The efficiency of statistical tools and a criterion for rejection of outlying observations," *Biometrica*, Vol. 28 (1936), pp. 308–320.

[7] D. A. DARLING, "The influence of the maximum term in the addition of independent random variables," *Trans. Am. Math. Soc.* (to be published).

[8] S. MALMQUIST, "On a property of order statistics from a rectangular distribution," *Skandinavisk Aktuarietidskrift*, Vol. 33 (1950), pp. 214–222.

[9] C. EISENHART, M. W. HASTAY AND W. A. WALLIS, "*Selected Techniques* of *Statistical Analysis*," McGraw-Hill Book Co., 1947.

# ON THE STOCHASTIC APPROXIMATION METHOD OF ROBBINS AND MONRO[1]

By J. Wolfowitz

*Cornell University*

**1. Summary.** In their interesting and pioneering paper Robbins and Monro [1] give a method for "solving stochastically" the equation in $x$: $M(x) = \alpha$, where $M(x)$ is the (unknown) expected value at level $x$ of the response to a certain experiment. They raise the question whether their results, which are contained in their Theorems 1 and 2, are valid under a condition (their condition (4′), our condition (1) below) which is statistically plausible and is weaker than the condition which they require to prove their results. In the present paper this question is answered in the affirmative. They also ask whether their conditions (33), (34), and (35) (our conditions (25), (26) and (27) below) can be replaced by their condition (5″) (our condition (28) below). A counterexample shows that this is impossible. However, it is possible to weaken conditions (25), (26) and (27) by replacing them by condition (3) (abc) below. Thus our results generalize those of [1]. The statistical significance of these results is described in [1].

**2. Statement of the problem.** Let $H(y \mid x)$ be a family of distribution functions which correspond to real values of the parameter $x$. Write

$$M(x) = \int_{-\infty}^{\infty} y \, dH(y \mid x).$$

We postulate that

$$(1) \qquad |M(x)| \leqq C < \infty, \qquad \int_{-\infty}^{\infty} (y - M(x))^2 \, dH(y \mid x) \leqq \sigma^2 < \infty,$$

and that either

$$(2) \qquad \begin{aligned} M(x) &\leqq \alpha - \delta, && \text{for } x < \theta, \\ M(x) &\geqq \alpha + \delta, && \text{for } x > \theta, \end{aligned}$$

for some $\delta > 0$, or else

$$(3a) \qquad \begin{aligned} &M(x) < \alpha \text{ for } x < \theta, \\ &M(\theta) = \alpha, \\ &M(x) > \alpha \text{ for } x > \theta, \end{aligned}$$

and, for some positive $\delta$,

(3b) $M(x)$ is strictly increasing if $|x - \theta| < \delta$

[1] Research under contract with the Office of Naval Research.

and

$$\inf_{|x-\theta|\geq\delta} |M(x) - \alpha| > 0. \tag{3c}$$

Let $\{a_n\}$ be a sequence of positive numbers such that

$$\sum_{n=1}^{\infty} a_n = \infty, \tag{4}$$

$$\sum_{n=1}^{\infty} a_n^2 < \infty. \tag{5}$$

Let $x_1$ be an arbitrary number. The Robbins-Monro convergence scheme is defined recursively for all $n$ by

$$x_{n+1} = x_n + a_n(\alpha - y_n),$$

where $y_n$ is a chance variable with distribution function $H(y \mid x_n)$. We shall prove the following

THEOREM. *If* (1), (4), (5), *and either* (2) *or* (3) (abc) *hold, then* $x_n$ *converges stochastically to* $\theta$.

**3. Proof of the Theorem.** Let

$$b_n = E(x_n - \theta)^2,$$

$$d_n = E[(x_n - \theta)(M(x_n) - \alpha)],$$

$$e_n = E\left[\int_{-\infty}^{\infty} (y - \alpha)^2 \, dH(y \mid x_n)\right].$$

Then, from (1),

$$0 \leqq e_n \leqq \sigma^2 + (C + |\alpha|)^2 = h \quad \text{(say)}. \tag{6}$$

An examination of the proof of [1] shows that, since (6) is valid, the following results of [1] hold under our assumptions:

$$b_n \to \text{a limit, say } b, \text{ as } n \to \infty; \tag{7}$$

$$d_n \geqq 0; \tag{8}$$

$$\sum_{n=1}^{\infty} a_n d_n < \infty; \tag{9}$$

$$b_{n+1} = b_1 + \sum_{1}^{n} a_j^2 e_j - 2 \sum_{1}^{n} a_j d_j. \tag{10}$$

From (4) and (9) we obtain

$$\liminf d_n = 0. \tag{11}$$

Let $n_1 < n_2 < \cdots$ be an infinite sequence of positive integers such that

$$\lim d_{n_j} = 0. \tag{12}$$

We assert that $x_{n_j}$ converges stochastically to $\theta$. If this were not so, there would exist an infinite subsequence $t_1 < t_2 < \cdots$ of the sequence of $n_j$, and positive numbers $\epsilon$ and $\eta$ such that, for all $j$,

$$P\{|x_{t_j} - \theta| > \eta\} > \epsilon. \tag{13}$$

But then for all $t_j$ we would have

$$\begin{aligned} d_{t_j} &= E[(x_{t_j} - \theta)(M(x_{t_j}) - \alpha)] \\ &= E[|x_{t_j} - \theta| \, |M(x_{t_j}) - \alpha|] \\ &\geqq \epsilon\eta \inf_{|x-\theta| \geqq \eta} |M(x) - \alpha|. \end{aligned} \tag{14}$$

From either (2) or (3) (abc) it follows that the last member of (14) is positive. This contradicts (12) and proves the assertion.

Let $\epsilon$ and $\eta$ be arbitrary positive numbers. Our theorem is proved if we can demonstrate the existence of an integer $N(\eta, \epsilon)$ such that, if $n > N(\eta, \epsilon)$,

$$P\{|x_n - \theta| > \eta\} \leqq \epsilon. \tag{15}$$

Let $s$ be a positive number such that

$$\frac{s^2 + s}{\eta^2} < \frac{\epsilon}{2}. \tag{16}$$

Since $x_{n_j}$ converges stochastically to $\theta$ there exists an integer $N_0$ such that

$$P\{|x_{N_0} - \theta| \geqq s\} < \frac{\epsilon}{2}, \tag{17}$$

and

$$\sum_{n=N_0}^{\infty} a_n^2 < \frac{s}{2h}. \tag{18}$$

Define, for $n > N_0$,

$$b'_n(z) = E\{(x_n - \theta)^2 \mid x_{N_0} = z\}.$$

From (6) and (10) we obtain

$$b'_n(z) \leqq (z - \theta)^2 + h \sum_{j=N_0}^{\infty} a_j^2 \leqq (z - \theta)^2 + \frac{s}{2}. \tag{19}$$

Consequently, when $n > N_0$,

$$P\left\{ |x_n - \theta| \geqq \eta \,\middle|\, |x_{N_0} - \theta| < s \right\} < \frac{s^2 + s}{\eta^2} < \frac{\epsilon}{2} \tag{20}$$

by (19) and (16). From (17) and (20) we conclude that (15) holds with $N(\eta, \epsilon) = N_0$. Thus our proof is complete.

**4. Two counterexamples.**

4.1. We show that (1) ((4′) in [1]) cannot be completely removed. Suppose that $\theta = \alpha = 0$, and define $H(y \mid x)$ for $x < 1$ in any way whatever provided only that

$$M(x) = x, \qquad -1 \leqq x < 1,$$
$$M(x) = -1, \qquad x < -1.$$

For all $x$ such that

$$\sum_{n=1}^{k} \frac{1}{n} \leqq x < \sum_{n=1}^{k+1} \frac{1}{n} \tag{21}$$

define $H(y \mid x)$ as follows:

$$H(y \mid x) = 0, \qquad y \leqq -1, \tag{22}$$

$$H(y \mid x) = 1 - \frac{1}{k^2}, \qquad -1 < y \leqq 2k^2 - 1, \tag{23}$$

$$H(y \mid x) = 1, \qquad y > 2k^2 - 1. \tag{24}$$

We have $M(x) \equiv 1$ for $x \geqq 1$. Thus $M(x)$ satisfies

$$\text{M}(x) \text{ is nondecreasing,} \tag{25}$$

$$M(\theta) = \alpha, \tag{26}$$

$$\left.\frac{dM(x)}{dx}\right|_{x=\theta} > 0. \tag{27}$$

(These are the conditions (33), (34), and (35) of [1].)

Now let $a_n = 1/n$ and $x_1 > 2$. We have

$$P\{x_{n+1} > x_n \mid y_1 = y_2 = \cdots = y_n = -1\} \geqq 1 - \frac{1}{n^2}$$

for all $n > 1$, and

$$P\{x_2 > x_1\} \geqq \tfrac{3}{4}.$$

Hence there is a positive probability that $x_n \to \infty$.

4.2. We show that the condition

$$\text{M}(x) < \alpha, \qquad x < \theta,$$
$$M(x) > \alpha, \qquad x > \theta, \tag{28}$$

(which is condition (5″) of [1]), and the condition (1) (or even the stronger condition (4) of [1], i.e., that there exist a positive constant $C'$ such that

$$\int_{-C'}^{C'} dH(y \mid x) = 1 \tag{29}$$

identically in $x$) are not sufficient for the theorem to hold. Let $a_n = 1/(4n)$, $\theta = \alpha = 0$. For all $x$ we define

$$H(y \mid x) = 0, \qquad y \leqq M(x), \tag{30}$$

$$H(y \mid x) = 1, \qquad y > M(x). \tag{31}$$

We define $M(0) = 0$, $M(1) = 1$. Let $x_1 = 1$. Then $x_2 = \frac{3}{4}$. If $x_n = y_n$ for $n \geqq 2$ we define $M(y_n) = (4n)^{-1}$. Since $\sum_{n=2}^{\infty} 1/(16n^2) < \frac{1}{2}$, we have

$$P\{x_n > \tfrac{1}{4}\} = 1$$

for all $n$. We can define $M(x)$ at points $x$ not included in our construction above in any manner compatible with (28).

The author is obliged to his colleague Professor J. C. Kiefer for helpful discussions while this paper was being written.

## REFERENCE

[1] Herbert Robbins and Sutton Monro, "A stochastic approximation method," *Annals of Math. Stat.*, Vol. 22 (1951), pp. 400–407.

# STOCHASTIC ESTIMATION OF THE MAXIMUM OF A REGRESSION FUNCTION[1]

By J. Kiefer and J. Wolfowitz

*Cornell University*

**1. Summary.** Let $M(x)$ be a regression function which has a maximum at the unknown point $\theta$. $M(x)$ is itself unknown to the statistician who, however, can take observations at any level $x$. This paper gives a scheme whereby, starting from an arbitrary point $x_1$, one obtains successively $x_2, x_3, \cdots$ such that $x_n$ converges to $\theta$ in probability as $n \to \infty$.

**2. Introduction.** Let $H(y \mid x)$ be a family of distribution functions which depend on a parameter $x$, and let

$$M(x) = \int_{-\infty}^{\infty} y \, dH(y \mid x). \tag{2.1}$$

We suppose that

$$\int_{-\infty}^{\infty} (y - M(x))^2 \, dH(y \mid x) \leqq S < \infty, \tag{2.2}$$

and that $M(x)$ is strictly increasing for $x < \theta$, and $M(x)$ is strictly decreasing for $x > \theta$. Let $\{a_n\}$ and $\{c_n\}$ be infinite sequences of positive numbers such that

$$c_n \to 0, \tag{2.3}$$

$$\sum a_n = \infty, \tag{2.4}$$

$$\sum a_n c_n < \infty, \tag{2.5}$$

$$\sum a_n^2 c_n^{-2} < \infty. \tag{2.6}$$

(For example, $a_n = n^{-1}$, $c_n = n^{-1/3}$.)

We can now describe a recursive scheme as follows. Let $z_1$ be an arbitrary number. For all positive integral $n$ we have

$$z_{n+1} = z_n + a_n \frac{(y_{2n} - y_{2n-1})}{c_n}, \tag{2.7}$$

where $y_{2n-1}$ and $y_{2n}$ are independent chance variables with respective distributions $H(y \mid z_n - c_n)$ and $H(y \mid z_n + c_n)$. Under regularity conditions on $M(x)$ which we shall state below we will prove that $z_n$ converges stochastically to $\theta$ (as $n \to \infty$).

The statistical importance of this problem is obvious and need not be discussed. The stimulus for this paper came from the interesting paper by Robbins and Monro [1] (see also Wolfowitz [2]).

[1] Research under contract with the Office of Naval Research. Presented to the American Mathematical Society at New York on April 25, 1952.

While we have no need to postulate the existence of the derivative of $M(x)$ (indeed, $M(x)$ can be discontinuous), the spirit of our regularity assumptions postulated below is as follows. (a) If $M(x)$ did have a derivative it would be zero at $x = \theta$. Hence we would have expected the derivative not to be too large in a neighborhood of $x = \theta$. (b) If, at a distance from $\theta$, $M(x)$ were very flat, then movement towards $\theta$ would be too slow. Hence outside of a neighborhood of $x = \theta$ we would have liked the absolute value of the derivative to be bounded below by a positive number. (c) If $M(x)$ rose too steeply in places we might through mischance get a movement of $z_n$ which would throw us far out from $\theta$. If there were many such steep places $z_n$ could be made to approach $+\infty$ or $-\infty$ with positive probability. We would therefore have postulated a Lipschitz condition.

From the mathematical point of view it would be aesthetic to weaken the conditions. From the practical point of view it might be objected that these conditions prevent $M(x)$ from being a function which flattens out toward the $x$-axis, for example, $M(x) = e^{-x^2}$, or from being a function which drops off steadily faster to $-\infty$, for example, $M(x) = -x^2$. Now in any practical situation one can always give a priori an interval $[C_1, C_2]$ such that $C_1 \leqq \theta \leqq C_2$. It will be sufficient if our conditions are fulfilled in this interval.

Suppose, however, that some $z_n \pm c_n$ falls outside the interval $[C_1, C_2]$ and one cannot take an observation at that level. If one then moves $z_n$ so that the offending $z_n \pm c_n$ is at $C_1$ or $C_2$, as the case may be, and proceeds as directed by (2.7), then our conclusion remains valid.

We postulate the following regularity conditions on $M(x)$.

CONDITION 1. There exist positive $\beta$ and $B$ such that

$$(2.8) \quad |x' - \theta| + |x'' - \theta| < \beta \text{ implies } |M(x') - M(x'')| < B|x' - x''|.$$

CONDITION 2. There exist positive $\rho$ and $R$ such that

$$(2.9) \quad |x' - x''| < \rho \text{ implies } |M(x') - M(x'')| < R.$$

CONDITION 3. For every $\delta > 0$ there exists a positive $\pi(\delta)$ such that

$$(2.10) \quad |z - \theta| > \delta \text{ implies } \inf_{\frac{1}{2}\delta > \epsilon > 0} \frac{|M(z+\epsilon) - M(z-\epsilon)|}{\epsilon} > \pi(\delta).$$

**3. Proof that $z_n$ converges stochastically to 0.** Let

$$(3.1) \quad b_n = E(z_n - \theta)^2,$$

$$(3.2) \quad U_n(z) = (z - \theta)\, E\{y_{2n} - y_{2n-1} \mid z_n = z\},$$

$$(3.3) \quad U_n^+(z) = \tfrac{1}{2}(U_n(z) + |U_n(z)|),\ U_n^-(z) = \tfrac{1}{2}(U_n(z) - |U_n(z)|),$$

$$(3.4) \quad P_n = E(U_n^+(z_n)),\ N_n = E(U_n^-(z_n)),$$

$$(3.5) \quad e_n = E(y_{2n} - y_{2n-1})^2.$$

From (2.7) we have

$$b_{n+1} = b_n + 2\frac{a_n}{c_n}(P_n + N_n) + \frac{a_n^2}{c_n^2}e_n. \tag{3.6}$$

Adding the expressions obtained from (3.6) for $b_{j+1} - b_j$ for $1 \leqq j \leqq n$, we obtain

$$b_{n+1} = b_1 + 2\sum_{j=1}^{n}\frac{a_j}{c_j}P_j + 2\sum_{j=1}^{n}\frac{a_j}{c_j}N_j + \sum_{j=1}^{n}\frac{a_j^2}{c_j^2}e_j. \tag{3.7}$$

Noting that $U_n^+(z) \geqq 0$ and that $U_n^+(z) > 0$ implies that $|z - \theta| < c_n$ because $M(x)$ is monotonic for $x < \theta$ and for $x > \theta$, it follows from (2.8) that, for all $n$ for which $c_n < \frac{1}{2}\beta$, we have

$$0 \leqq U_n^+(z) < 2Bc_n^2 \tag{3.8}$$

It follows from (2.5) and (3.8) that the positive-term series

$$\sum_{n=1}^{\infty}\frac{a_n}{c_n}P_n \tag{3.9}$$

converges, say to $\alpha$. From (2.9) we have

$$[M(z_n + c_n) - M(z_n - c_n)]^2 < R^2 \tag{3.10}$$

for $n$ sufficiently large. Also for large enough $n$,

$$\begin{aligned} &E\{(y_{2n} - y_{2n-1})^2 \mid z_n\} \\ &= E\{(y_{2n} - M(z_n + c_n))^2 + (y_{2n-1} - M(z_n - c_n))^2 \mid z_n\} \\ &\qquad + [M(z_n + c_n) - M(z_n - c_n)]^2 \leqq 2S + R^2 \end{aligned} \tag{3.11}$$

by (2.2) and (3.10). Hence for large enough $n$

$$E[y_{2n} - y_{2n-1}]^2 \leqq 2S + R^2. \tag{3.12}$$

Consequently from (2.6) we obtain that the positive-term series

$$\sum_{n=1}^{\infty}\frac{a_n^2}{c_n^2}e_n \tag{3.13}$$

converges, say to $\gamma$. Hence, since $b_{n+1} \geqq 0$, it follows from (3.7) that

$$2\sum_{j=1}^{n}\frac{a_j}{c_j}N_j \geqq -b_1 - 2\alpha - \gamma > -\infty, \tag{3.14}$$

so that the negative-term series

$$\sum_{n=1}^{\infty}\frac{a_n}{c_n}N_n \tag{3.15}$$

converges.

Let

$$K_n = \left| \frac{M(z_n + c_n) - M(z_n - c_n)}{c_n} \right|. \tag{3.16}$$

Then

$$E\{K_n \mid z_n - \theta \mid\} = \frac{P_n - N_n}{c_n}. \tag{3.17}$$

From the convergence of (3.9) and (3.15) and the divergence of $\sum a_n$, it follows that

$$\liminf_{n \to \infty} E\{K_n \mid z_n - \theta \mid\} = 0. \tag{3.18}$$

Let $n_1 < n_2 < \cdots$ be an infinite sequence of positive integers such that

$$\lim_{j \to \infty} E\{K_{n_j} \mid z_{n_j} - \theta \mid\} = 0. \tag{3.19}$$

We assert that $(z_{n_j} - \theta)$ converges stochastically to zero as $j \to \infty$. For if not, there would exist two positive numbers $\delta$ and $\epsilon$ and a subsequence $\{t_j\}$ of $\{n_j\}$ such that, for all $j$,

$$P\{\mid z_{t_j} - \theta \mid > \delta\} > \epsilon, \tag{3.20}$$

which implies that

$$E\{K_{t_j} \mid z_{t_j} - \theta \mid\} \geqq \delta\epsilon\pi\left(\frac{\delta}{2}\right) > 0 \tag{3.21}$$

for all $j$ for which $c_{t_j} < \frac{1}{2}\delta$. But (3.21) contradicts (3.19) and the stochastic convergence to zero of $(z_{n_j} - \theta)$ is proved.

Let $\eta$ and $\epsilon$ be arbitrary positive numbers. The proof of the theorem will be complete if we can show the existence of an integer $N(\eta, \epsilon)$ such that

$$P\{\mid z_n - \theta \mid > \eta\} \leqq \epsilon \text{ for } n > N(\eta, \epsilon). \tag{3.22}$$

Let $s$ be a positive number such that

$$\frac{s^2 + s}{\eta^2} < \frac{\epsilon}{2}. \tag{3.23}$$

Because $z_{n_j}$ converges stochastically to $\theta$ there exists an integer $N_0$ such that

$$P\{\mid z_{N_0} - \theta \mid \geqq s\} < \frac{\epsilon}{2}. \tag{3.24}$$

We may also choose $N_0$ so large that

$$c_n < \min\left(\frac{\rho}{2}, \frac{\beta}{2}\right) \text{ for all } n \geqq N_0, \tag{3.25}$$

and

$$\sum_{n=N_0}^{\infty} \frac{a_n^2}{c_n^2} < \frac{\epsilon}{2R^2 + 4S}, \tag{3.26}$$

and

$$\sum_{n=N_0}^{\infty} a_n c_n < \frac{\epsilon}{8B}. \tag{3.27}$$

Proceeding in a manner similar to that used to obtain (3.7), we have, for each $n > N_0$,

$$\begin{aligned} E\{(z_n - \theta)^2 \mid z_{N_0} = z\} &= (z - \theta)^2 + 2 \sum_{j=N_0}^{n-1} \frac{a_j}{c_j} E\{U_j \mid z_{N_0} = z\} \\ &+ \sum_{j=N_0}^{n-1} \frac{a_j^2}{c_j^2} E\{(y_{2j} - y_{2j-1})^2 \mid z_{N_0} = z\} \end{aligned} \tag{3.28}$$

$$\leqq (z - \theta)^2 + 2 \sum_{j=N_0}^{\infty} \frac{a_j}{c_j} E\{U_j^+ \mid z_{N_0} = z\} + (R^2 + 2S) \sum_{j=N_0}^{\infty} \frac{a_j^2}{c_j^2} < (z - \theta)^2 + \epsilon.$$

Using (3.23), (3.28), and Tchebycheff's inequality, we have

$$P\{|z_n - \theta| > \eta \,\Big|\, |z_{N_0} - \theta| < \epsilon\} < \frac{\epsilon}{2}. \tag{3.29}$$

The inequalities (3.24) and (3.29) show that (3.22) holds for $N(\eta, \epsilon) = N_0$, and the proof is complete.

**4. Further problems.** The following remarks about further problems apply also to [1].

*A*. An obvious problem is to determine sequences $\{c_n\}$ and $\{a_n\}$ which would be optimal in some reasonable sense.

*B*. An important problem is to determine a stopping rule, that is, a rule by which the statistician decides when he is sufficiently close to $\theta$.

*C*. This problem is a combination of *B* and a generalization of *A*, that is, to determine an optimal procedure with its stopping rule.

## REFERENCES

[1] H. Robbins and S. Monro, "A stochastic approximation method," *Annals of Math. Stat.*, Vol. 22 (1951), pp. 400–407.

[2] J. Wolfowitz, "On the stochastic approximation method of Robbins and Monro," *Annals of Math. Stat.*, Vol. 23 (1952), pp. 457–461.

# NOTES

## AN ASYMMETRIC BELL-SHAPED FREQUENCY CURVE

By Fausto I. Toranzos

*Mendoza, Argentina*

**1. Summary.** Frequency functions of the form (1) below including Pearson's system as a special case are derived from a generating differential equation (2). Their properties are discussed, methods of fitting them are suggested, and their characteristic function is determined.

**2. Introduction.** We wish to contribute to the study of the frequency functions of the form[1]

$$y = f(x) \exp\{-\tfrac{1}{2}a^2 (x - b)^2\}. \tag{1}$$

We will realize this purpose by taking as a generating differential equation of our system of frequency functions

$$\frac{y'}{y} = \frac{Q_{m+1}(x)}{P_m(x)} \tag{2}$$

in which $Q$ is a polynomial of degree $m + 1$ and $P$ another of degree $m$. Carrying out the division of the polynomials of (2) gives

$$\frac{y'}{y} = \alpha x + \beta + \frac{Q_{m-1}(x)}{P_m(x)} \tag{3}$$

and integrating

$$y = k \exp\{\tfrac{1}{2}\alpha x^2 + \beta x + \varphi(x)\}, \qquad \varphi(x) = \int \frac{Q_{m-1}(x)}{P_m(x)}\, dx.$$

The integral $\varphi(x)$ can always be evaluated by decomposing it into integrals of the forms

$$\int \frac{dx}{(x - a)^r}, \qquad \int \frac{Mx + N}{(x^2 + px + q)^s}\, dx.$$

Bearing this in mind, it is easy to show that, if $\alpha < 0$ and the roots of the denominator are simple, $y$ is equal to the Gauss function multiplied by Pearson's functions. When there are multiple roots, exponential factors will appear.

Pearson's system is a particular case of (3) when $\alpha = 0$, $\beta = 0$ and $m = 2$.

[1] In this work we generalize and complete the results that we have given in [1].

**3. The case $m = 1$.** The case $m = 1$ is particularly important. The corresponding differential equation is

$$\frac{y'}{y} = \frac{A_2x^2 + A_1x + A_0}{B_1x + B_0}.$$

Making a change in the origin of $x$, we have

$$\frac{y'}{y} = \frac{a_2x^2 + a_1x + a_0}{x}, \tag{4}$$

and integrating,

$$y = k \exp\left\{\left(x + \frac{a_1}{a_2}\right)^2\right\} x^{a_0}.$$

This function gives bell-shaped curves when $a_2 < 0$, $a_0 > 0$. Now putting $a_2 = -a^2$, $a_1/a_2 = b$, $a_0 = c$, we have

$$y = kx^c \exp\{-\tfrac{1}{2}a^2(x - b)^2\}. \tag{5}$$

The graph of $y$ is a bell-shaped curve with zeros at the points $x_1 = 0$, $x_2 = \infty$, if we take as the range of $x$, $0 \leqq x \leqq \infty$. In this interval $y$ has only one maximum.

This case if of great importance due to its possible applications in distributions of bell-shaped frequencies with a finite range to the left of the maximum and extending to infinity on the right. Such happens in the case of the distribution of relative percentage values of prices and volume, in which the decrease is limited (from 100% to 0%), whereas the increase is theoretically unlimited (from 100% on). In spite of the natural asymmetry, the curve is remarkably close to the Gauss curve in the vicinity of the maximum.

When $a_2 > 0$ and $a_0 < 0$ we have a U-shaped curve which has a minimum near $x = b$, and which becomes infinite at $x = 0$ and $x = \infty$.

When $a_2 < 0$ and $a_0 < 0$ we have a curve which has a zero at the point $x = \infty$, a minimum near the origin, and a maximum near $x = b$, and which becomes infinite at the origin.

When $a_2 > 0$ and $a_0 > 0$ the curve has a zero at the point $x = 0$, a maximum near $x = 0$, and a minimum near $x = b$, and becomes infinite at $x = \infty$.

**4. Determination of the constants.** In order to fit the function (5) to an empirically determined frequency series, we may use the method of moments in the same way as Pearson does for his system.

The differential equation (2) can be expressed as

$$xdy = y(a_2x^2 + a_1x + a_0).$$

Multiplying by $x^s$ and integrating from 0 to $\infty$, we have

$$-(s + 1)m_s = a_2m_{s+2} + a_1m_{s+1} + a_0m_s,$$

in which $m_s$ are moments of order s. Identifying the functional moments with the empirical moments and giving successively the values 0, 1, and 2 to $s$, we

have a system of three equations with three unknowns from which we can determine $a_2$, $a_1$, $a_0$, and therefore $a$, $b$, and $c$.

We can also determine the constants by the least squares method using the equation

$$\ln y = \tfrac{1}{2}a_2 x^2 + a_1 x + a_0 \ln x + C.$$

The value of $C$ should be adjusted so that it satisfies the condition

$$\int_0^\infty y\,dx = 1.$$

**5. Determination of the characteristic function.** The characteristic function $\phi(t)$ of (5) is defined by

$$\phi(t) = k\int_0^\infty x^c \exp\{-\tfrac{1}{2}a^2(x-b)^2 + ixt\}\,dx.$$

Multiplying out the square, and putting $ab = p$, $ax = u$ we have

$$\phi(t) = ka^{-(c+1)} \exp\{-\tfrac{1}{2}a^2b^2\} \int_0^\infty u^c \exp\left\{-\tfrac{1}{2}u^2 + \left(p + \frac{it}{a}\right)u\right\} du$$

$$(6) \qquad = ka^{-(c+1)} \exp\{-\tfrac{1}{2}a^2b^2\}\, L_{-(p+it/a)}[u^c \exp\{-\tfrac{1}{2}u^2\}]$$

$$= ka^{-(c+1)}\,\Gamma(c+1) \exp\{-\tfrac{1}{2}a^2b^2 - \tfrac{1}{4}(p + it/a)^2\}\, D_{-(c+1)}[-p - it/a],$$

writing $L$ for the Laplace transform, and $D_c(z)$ for Weber's function of the parabolic cylinder [2].

This function can be expressed by Whittaker's confluent hypergeometric function, which has been tabulated [2].

To determine the value of $K$ it is sufficient to make $t = 0$ in (6), giving

$$\phi(0) = 1 = ka^{-(c+1)}\,\Gamma(c+1) \exp\left\{-\tfrac{1}{2}a^2b^2 - \frac{p^2}{4}\right\} D_{-(c+1)}[-p],$$

and therefore

$$k = \frac{a^{(c+1)}}{\Gamma(c+1)} \exp\left\{\tfrac{1}{2}a^2b^2 + \frac{p^2}{4}\right\} \Big/ D_{-(c+1)}[-ab],$$

and finally

$$\phi(t) = \exp\left\{\tfrac{1}{2}ibt - \frac{t^2}{a^2}\right\} D_{-(c+1)}[-ab - it/a]/D_{-(c+1)}[-ab].$$

Now we can determine the moments and other characteristics of the function (5).

## REFERENCES

[1] Fausto I. Toranzos, "A system of frequency curves which generalizes that of Pearson," *Revista Fac. Ci. Econ. Univ. Cuyo*, Vol. 1 (1949), 7 pp. (Spanish).

[2] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, Cambridge University Press, 1927, pp. 347-349.

# REMARKS ON A MULTIVARIATE TRANSFORMATION[1]

By Murray Rosenblatt

*University of Chicago*

The object of this note is to point out and discuss a simple transformation[2] of an absolutely continuous $k$-variate distribution $F(x_1, \cdots, x_k)$ into the uniform distribution on the $k$-dimensional hypercube. A discussion of related transformations has been given by P. Lévy [1].

Let $X = (X_1, \cdots, X_k)$ be a random vector with distribution function $F(x_1, \cdots, x_k)$. Let $z = (z_1, \cdots, z_k) = Tx = T(x_1, \cdots, x_k)$, where $T$ is the transformation considered. Then $T$ is given by

$$z_1 = P\{X_1 \leqq x_1\} = F_1(x_1),$$

$$z_2 = P\{X_2 \leqq x_2 \mid X_1 = x_1\} = F_2(x_2 \mid x_1),$$

$$\vdots$$

$$z_k = P\{X_k \leqq x_k \mid x_{k-1} = x_{k-1}, \cdots, X_1 = x_1\} = F_k(x_k \mid x_{k-1}, \cdots, x_1).$$

One can readily show that the random vector $Z = TX$ is uniformly distributed on the $k$-dimensional hypercube, for

$$P\{Z_i \leqq z_i; i = 1, \cdots, k\}$$

$$= \int_{\{z \mid z_i \leqq z_i\}} \cdots \int d_{x_k} F_k(x_k \mid x_{k-1}, \cdots, x_1) \cdots d_{x_1} F_1(x_1)$$

$$= \int_0^{z_k} \cdots \int_0^{z_1} dz_1 \cdots dz_k = \prod_{i=1}^{k} z_i,$$

when $0 \leqq z_i \leqq 1$, $i = 1, \cdots, k$. Hence $Z_1, \cdots, Z_k$ are uniformly and independently distributed on [0, 1].

Let $X_{(i)} = (X_{1(i)}, \cdots, X_{k(i)})$, $i = 1, \cdots, n$, be a random sample of $n$ vectors from a population with distribution function $F(x_1, \cdots, x_k)$ and let $G(x_1, \cdots, x_k)$ be the corresponding sample distribution function. It has been found that the probability distribution of the Kolmogorov-Smirnov statistic,

$$\max_{x_1, \cdots, x_k} | F(x_1, \cdots, x_k) - G(x_1, \cdots, x_k) |,$$

is not the same for all continuous $F$ when $k > 1$ [2]. The same can be said for the multidimensional von Mises statistic

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (F(x_1, \cdots, x_k) - G(x_1, \cdots, x_k))^2 \, d_{x_1 \cdots x_k} F(x_1, \cdots, x_k).$$

However, it would still be of interest to study the Kolmogorov-Smirnov and von Mises statistics for the case of sampling from a population uniformly distributed on the $k$-dimensional hypercube. One could test whether the $X_{(i)}$, $i = 1, \cdots, n$, are a sample from a population with distribution function $F(x_1, \cdots, x_k)$ by

[1] Work done under ONR contract.

[2] I have recently learned that J. H. Curtiss and I. R. Savage have also considered this transformation.

using either the Kolmogorov-Smirnov or von Mises statistic to test whether the $Z_{(i)} = TX_{(i)}$ are a sample from a population uniformly distributed on the $k$-dimensional hypercube. Both statistics are invariant under permutation of the $X_{(i)}, i = 1, \cdots, n$.

The transformation is of some interest in setting up regions of equal probability mass in a chi-square test of goodness of fit. Assume again that one wishes to test whether the $X_{(i)}, i = 1, \cdots, n$, come from a population with absolutely continuous distribution function $F(x_1, \cdots, x_k)$. This, as remarked before, is equivalent to testing whether the $Z_{(i)} = TX_{(i)}, i = 1, \cdots, n$, come from a population with distribution function $z_1 \cdots z_k$, where $0 \leqq z_j \leqq 1, j = 1, \cdots, k$. Divide the $k$-dimensional hypercube into the $N^k$ cells

$$C_{j_1 \cdots j_k} = \left\{ z \,\middle|\, \frac{j_i}{N} < z_i < \frac{j_i + 1}{N},\ i = 1, \cdots, k \right\},$$

where $j_1, \cdots, j_k = 0, 1, \cdots, N - 1$. Each of these cells has probability mass $1/N^k$ under the null hypothesis. Let $\nu_{j_1 \cdots j_k}$ be the number of transformed observations $Z_{(i)} = TX_{(i)}$ in $C_{j_1 \cdots j_k}$. The chi-square statistic

$$\sum \frac{\left( \nu_{j_1 \cdots j_k} - \dfrac{n}{N^k} \right)^2}{n/N^k}$$

can be used to test whether the sample $X_{(i)}$ comes from a population with distribution function $F(x_1, \cdots, x_k)$. A test for goodness of fit based on the chi-square statistic above is consistent as $n, N \to \infty$.

There are $k!$ transformations $T$ of the type described above corresponding to the $k!$ ways in which one can number the coordinates $x_1, \cdots, x_k$. One might think this unsatisfactory because of a suspicion that the experimenter might carry out one of the procedures described above with each of the $k!$ transformations and then choose that particular transformation which yields the result he wishes to obtain. But this situation can arise in any case where there is a multitude of tests in the same context.

The transformation $T$ can be written down explicitly in several cases. We shall write down the transformation when $F(x_1, \cdots, x_k)$ is a normal distribution with mean $M = (m_1, \cdots, m_k)$ and covariance matrix $\Lambda = \{ \lambda_{ij} \}, i, j = 1, \cdots, k$. Let $\Lambda^{(r)} = \{ \lambda_{ij} \}, i, j = 1, \cdots, r \leqq k$, and $\Lambda_{ij}^{(r)}$ be the cofactor of $\{ \lambda_{ij} \}$ in $\Lambda^{(r)}$. The transformation $T$ is then given by

$$F_1(x_1) = \Phi\left( \frac{x_1 - m_1}{\sqrt{\lambda_{11}}} \right),$$

$$F_2(x_2 \mid x_1) = \Phi\left( \frac{x_2 - m_2 + (\Lambda_{21}^{(2)}/\Lambda_{22}^{(2)})(x_1 - m_1)}{\sqrt{\Lambda^{(2)}/\Lambda_{22}^{(2)}}} \right),$$

$$\vdots$$

$$F_k(x_k \mid x_{k-1}, \cdots, x_1) = \Phi\left( \frac{x_k - m_k + \sum_{j=1}^{k-1} (\Lambda_{kj}/\Lambda_{kk})(x_j - m_j)}{\sqrt{\Lambda/\Lambda_{kk}}} \right).$$

Let $F(x_1, x_2)$ be a normal distribution with means $m_1$, $m_2$, variances $\sigma_1^2$, $\sigma_2^2$ and correlation coefficient $\rho$. The transformation can then be written as

$$F_1(x_1) = \Phi\left(\frac{x_1 - m_1}{\sigma_1}\right),$$

$$F_2(x_2 \mid x_1) = \Phi\left(\frac{x_2 - m_2 + \frac{\rho\sigma_1}{\sigma_2}(x_1 - m_1)}{\sigma_2\sqrt{1 - \rho^2}}\right).$$

REFERENCES

[1] P. Lévy, *Théorie de l'Addition des Variables Aléatoires*, Gauthier-Villars, Paris, 1937, pp. 71–73, 121–123.

[2] Paul B. Simpson, "Note on the estimation of a bivariate distribution function," *Annals of Math. Stat.*, Vol. 22 (1951), pp. 476–478.

## ABSTRACTS OF PAPERS

*(Abstracts of papers presented at the Eugene meeting of the Institute, June 19–21, 1952)*

1. **The Auditory Cortex—A Probability Model.** Archie R. Tunturi, University of Oregon Medical School.

The role played by the brain in communication is well known, but in what manner the brain handles information is not understood. Some progress has been made in this direction by studying the anatomy and physiology of the auditory cortex in the anesthetized dog with controlled acoustic signals. Communication may be thought of as making a representation in a space of a representation in another space. In three of the four auditory areas (on one side of the brain), the entire frequency spectrum from 100 to 12800 cps is represented literally spacewise by groups of cells that respond only to a narrow range of frequencies. A special method increases the signal to noise ratio, by augmenting the electrical response of the cells, thereby permitting exact measurements of the characteristic frequency and intensity for each group of cells. This is similar to a narrow band filter, and does not reveal the effect of other frequencies on the information. The information capacity of the system can be inferred if it can be assumed that occurrence of the augmented response for the group of cells follows some probability function. These probabilities for all groups of cells can be assembled into a model representing the behavior of the system as a communication device. If there are 70 groups of cells between 100 and 12800 cps, the probability of any particular combination would be $1/2^{70}$, if the selections were equally probable. The effect of noise on this system will be considered. (Research sponsored in part by the Office of Naval Research.)

2. **Testing Message Diffusion: The Utilization of Mathematical Models.** Stuart C. Dodd, Richard J. Hill, and Susan Huffaker, University of Washington.

In connection with the study of interpersonal verbal communication, the Washington Public Opinion Laboratory designed an experimental procedure which yielded data on the temporal diffusion of thirty-three different messages in a population of 184 individuals. Data (including the recipient of each message, the initiator of communication, and the time of communication) were obtained on 5,522 separate instances of communication. The

experimental population was selected so that the assumption could reasonably be made that it was homogeneous relative to the interacting (i.e., each individual had an equal probability of interacting), and uninfluenced by the spatial factor in diffusion. In the attempt to describe the temporal diffusion of the messages, three rational mathematical models were compared with the empirical results. The models utilized were the logistic, the normal ogive, and the binomial distribution, for which the mathematical conditions are well identified. The experiment had been designed to match those social conditions underlying the communication with the mathematical conditions underlying the logistic model (namely, a constant probability of interacting in the population and through time). The higher and lower degrees of descriptive closeness and goodness of fit for the three models, and for ordinal and cardinal time units separately, are reported and discussed.

### 3. On the Method of Collective Marks. HERMAN RUBIN, Stanford University.

In *Colloques International du Centre National de la Recherche Scientifique, No. 13*, D. van Dantzig considered the application of the method of generating functions to run problems. Let $x_1, \cdots, x_n, \cdots$ be a sequence of nonnegative integer-valued random variables of which only a finite number are nonzero with probability one. Then $E(\Pi_{i=1}^{\infty} t_i^{x_i})$ exists for $|t_i| \leq 1$. If $0 \leq t_i \leq 1$, then we can interpret $t_i$ as the probability that a coin when tossed will come up heads. Let an experiment be performable independently an arbitrary number of times with results $a_j$ with probability $p_j$, and let us ask questions concerning the numbers of runs of various types obtained. One may compute the generating function of the various quantities involved by performing the experiment $n$ times with probability $\lambda^n(1-\lambda)$. If $y_j$ is the number of occurrences of $a_j$ and $z_{jk}$ is the number of runs of length $k$ of type $j$, then $P = E(S^n \Pi u_j^{y_j} \Pi v_{jk}^{z_{jk}}) = 1 - \lambda + \Sigma_j f_j P_j$, where $f_j = \Sigma_k (\lambda p_j S u_j)^k v_{jk}$ and $P$ is the probability that no tail occurs and $P_j$ is the probability that in addition $a_j$ is not the first outcome of the experiment. Then $P_j = P - f_j P_j$ and hence $P = (1-\lambda)/(1 - \Sigma_j f_j/(1+f_j))$. Then such questions as to number and type of runs in $n$ experiments can be answered from the conditional generating function. Other types of questions can be answered by letting $\lambda$ approach 1.

### 4. On Asymptotic Properties of Estimates. LUCIEN LECAM, University of California, Berkeley.

Let $f(x, \theta)$ be a probability density depending on a parameter $\theta$. Under regularity conditions on $f(x, \theta)$ it is shown that, for a broad class of loss functions and for positive continuous a priori densities: (1) The maximum liklihood (M. L.) estimate is asymptotically equivalent to a Bayes estimate. Moreover all "regular" asymptotic Bayes estimates are equivalent. (2) Given any estimate $T_n$ there exists a randomized estimate $T_n'$ asymptotically equivalent to $T_n$ and depending on the sample point by means of the M. L. estimate only. (3) If an estimate $T_n$ is super-efficient at a point $\theta_0$ there exists a sequence $\theta_n$, $\theta_n \to \theta_0$ of points at which $T_n$ is worse than the M. L. estimate, except maybe for a finite number of values of $n$.

### 5. The Principle of Invariance Applied to a Problem of Classification. ROSEDITH SITGREAVES, Stanford University.

A problem of classification considered by Wald and Anderson is as follows: an individual is to be classified on the basis of $p$ measurements as coming from one of two populations, $\Pi_1$ or $\Pi_2$. The distribution of the measurements in each population is multivariate normal with the same covariance matrix, but with different expected values. The parameters of the distribution are unknown, but samples of $N_1$ observations from $\Pi_1$ and $N_2$ observations from $\Pi_2$ are available. Both Wald and Anderson proposed classification procedures determined by substituting sample estimates for unknown parameter values in the ap-

propriate Bayes' solution for the case when the parameters are known. Optimum properties for these procedures are difficult to prove because the two alternatives are composite. If the possible procedures are restricted by adopting the principle of invariance, there is obtained a family of classification procedures depending on a single nuisance parameter such that the class of procedures defined for a fixed value of the nuisance parameter is a complete class. In a single case, namely, when $N_1 = N_2$ and the losses due to misclassification are equal, the minimax procedure is independent of the nuisance parameter, and depends upon classification statistics proposed by Wald and Anderson.

6. **Optimum Selection Procedures.** EDWARD PAULSON, University of Washington.

Let $(Y_1, Y_2, \cdots, Y_k, X)$ be a $(k + 1)$ dimensional random variable, with $Y_1, Y_2, \cdots, Y_k$ denoting the admission scores of an individual on a series of $k$ admission tests and $X$ the performance or achievement score of the same individual. Optimum procedures for selecting or rejecting an individual on the basis of his observed admission scores have been investigated by Birnbaum, Chapman, and Cochran. By considering the $X$ scores as playing the role of a parameter with respect to the conditional distribution of $Y_1, Y_2, \cdots, Y_k$ when $X$ is fixed, two additional procedures are derived, each optimum in a certain well defined sense. The second of these procedures, which is analogous to a Bayes solution and minimizes the average loss in classifying an individual, can easily be extended to the problem of classifying an individual on the basis of his admission scores into one of 3 or more classes, providing that the appropriate loss functions and the conditional probability distribution of $X$ for fixed values of $Y_1, Y_2, \cdots, Y_k$ are known.

7. **A Theorem on Convex Cones with Applications to Linear Inequalities.** JERRY W. GADDUM, Institute for Numerical Analysis.

I can summarize the results as follows: It is proved that if $A$ is a convex cone in $E_n$ and $A^*$ its polar cone, $A$ and $A^*$ have a non-null vector in common unless $A$ is an $E_r$. This is applied to obtain the result that $Ax \geq 0$ has a solution if and only if $AA'y \geq 0$ has a non-negative solution. Results due to Dines and to Motzkin are then applicable to obtain further existence theorems. The theorem on convex cones also provides a method of solution of systems of inequalities.

8. **On the Prediction of Nonstationary Stochastic Processes.** R. C. DAVIS, Bureau of Ordnance.

During a finite time interval $T$ the real valued function $S(t) + N(t)$ is observed, in which $S(t)$ is a signal and $N(t)$ is a linearly superimposed noise disturbance. The problem is to predict the value of a given linear functional of $S(t)$, the predictor formula having certain preassigned "optimum properties" among a certain class of predictors. In the case in which the mean value of $S(t)$ is known, the random components of $S(t)$ and $N(t)$ are strictly stationary, and the time interval $T$ is infinite, a complete solution to this problem has been given by N. Wiener. (In the case of discrete time series, the solution was given by A. Kolmogorov.) This theory has been extended by L. Zadeh and J. Ragazzini to the case in which $T$ is a finite time interval and the mean value of $S(t)$ is unknown but is restricted to be a polynomial in time. The above theory is extended to the case in which the random components of both $S(t)$ and $N(t)$ are nonstationary in time and merely possess finite continuous covariance and cross covariance functions. The analytical methods of probability theory used are those developed by M. Loève and K. Karhunen in their studies of stochastic processes of second order. These techniques are very powerful in the analysis of transient random phenomena in linear systems. Finally an exposition is given of prediction by the method of conditional probabilities, and this method is compared with the least squares method.

**9. Power Efficiency Function for Normal Alternatives for Several Nonparametric Tests.** W. J. DIXON, University of Oregon.

Power curves are computed for samples of size 3, 4, 5 for the maximum absolute deviation, the rank-sum, and the median tests for normal alternatives. These curves are compared with the power curves for the $t$-test. (Research sponsored by the Office of Naval Research.)

**10. Bounds for Second Moments of the Sample Range.** SIGEITI MORIGUTI, University of North Carolina and University of Tokyo.

Let $R$ be the range, that is the difference between the two extreme values, of a sample of size $n$ from a symmetrical population with variance $\sigma^2$. This paper deals with the least upper bound for $E(R^2)/\sigma^2$, the greatest lower bound for the coefficient of variation of $R$, and the greatest lower bound for $V(R)/\sigma^2$. It is shown that these can be obtained as characteristic values of certain integral equations. For $n$ even, these equations can be reduced to linear algebraic equations. Numerical results are given for small values of $n$.

**11. Estimates, Tests, and Tolerance Intervals for Lump Data from a Normal Distribution.** HARRY M. HUGHES, University of California, Berkeley.

If $N$ independent samples from a normal distribution with unknown mean $\mu$ and unknown variance $\sigma^2$ are lumped into $m$ subsamples of known size, and if the sums of the subsample values are observed but not the individual values, then the variance of the maximum likelihood estimate of $\mu$ is unchanged but the variance of the maximum likelihood estimate of $\sigma^2$ (modified to remove bias) is increased to that of a sample of size $m$. The effect on confidence intervals and on standard tests such as the $t$-test and $F$-test follows immediately.

The effect on tolerance intervals, not heretofore easily available, is presented in the form of a table of factors, depending on $m$, $N$ and the confidence coefficient $\gamma$ by which to modify existing tables of the coefficient $\lambda$ in the expression $\bar{x} \pm \lambda s$. The range covered is $m = 2$ to 100, $N = 3$ to 1000 for $N > m$; confidence coefficients .75, .90 and .99; and all tolerance coefficients (percentage of population to be covered) since table entries are independent of the tolerance coefficient. The resulting increase in the expected length of the tolerance interval is indicated.

**12. A Computing Formula for the Power of the Two-Sided t-Test for Even Degrees of Freedom.** WESLEY L. NICHOLSON, University of Oregon.

A general formula is exhibited for the power function of the "two-sided" $t$-test for even degrees of freedom. The form employed is particularly adapted to computation of the power for various alternative means, given fixed arbitrary type I error and even degrees of freedom. The power functions for 2, 4, 6, 8, and 10 degrees of freedom are written out explicitly with an indication of their use for arbitrary type I error.

**13. On the Analysis of Data Matched in Pairs.** FRANK J. MASSEY, JR., University of Oregon.

We suppose that groups of four observations are drawn under varying environmental conditions for the different groups. In each group two of the individuals are subjected to one treatment and the other two are subjected to a second treatment. This paper discusses various parametric and nonparametric techniques of analyzing the data. Power curves are given for the case where the two populations are normally distributed differing only in mean value.

**14. The Power Function of the Haldane-Smith Test.** B. M. BENNETT, University of Washington.

Haldane and Smith (*Annals of Eugenics*, 1948) have devised a non-parametric test for the existence of a birth order effect in families in which a certain number of sibs are affected by a particular disease. The test consists of summing the ranks of all affected individuals. The distribution has been worked out under the assumption that all ranks are equally likely. It is shown that this test may be regarded more generally as a test for the constancy of the probability of an event from trial to trial when it is known to occur a fixed number of times out of a given number of trials. The power function of the test has been considered with respect to several classes of alternatives concerning the successive probabilities of the event.

**15. Spectral Analysis of Stationary Time Series. (Preliminary Report.)** ULF GRENANDER AND MURRAY ROSENBLATT, University of Chicago.

Having observed a sample sequence (finite) of a vector-valued stationary time series, one wishes to make inferential statements about the spectrum of the underlying process on the basis of the sample. Assume that the process is a linear process and that the spectrum satisfies certain regularity conditions. There are three aspects to the work being carried out. 1. Obtaining curve-estimates of the spectral density and spectral distribution function. These estimates have been chosen as different types of quadratic forms in the sample values. 2. Obtaining confidence bands and confidence regions for both the spectral density and spectral distribution function. Certain limit theorems nonparametric with respect to the spectrum have been obtained in connection with this aspect of the problem. 3. These limit theorems have also been used to construct one and two sample tests for time series. We plan to set up a sampling program to test the approach to the limiting distributions. Artificial time series will be generated and submitted to analytic techniques provided by the theory. In this connection, schemes of computation are now being devised. Auxiliary tables are under preparation. (This work is being done under the sponsorship of the Office of Naval Research.)

**16. Asymptotically Subminimax and Asymptotically Admissible Statistical Decision Procedures.** JACK LADERMAN, Columbia University.

A study is made of the asymptotic behavior of the minimax decision procedure when the distribution space, $\Omega$, consists of a finite number of absolutely continuous distribution functions and the decision problem is to select the true distribution function when a sample of size $n$ is known. It is seen that "asymptotically subminimax" procedures frequently exist which are clearly more desirable than the minimax procedure when $n$ is sufficiently large. Let $r_i(T_n)$ denote the risk associated with the decision procedure $T_n$ when the $i$-th distribution is the true distribution. A sequence of decision procedures, $\{T_n\}$, is said to be asymptotically admissible if there does not exist another sequence of decision procedures, $\{T_n'\}$, such that $\overline{\lim}_{n\to\infty} r_i(T_n')/r_i(T_n) \leqq 1$ for all values of $i$ and the strict inequality holds for at least one value of $i$. A detailed study is made of the class of problems for which $\Omega$ consists of $k$ univariate normal distributions all having the same variance but different means. The minimax decision procedure is found to be asymptotically admissible only for very special sets of values for the $k$ means. For other sets of mean values, asymptotically subminimax procedures are obtained which are also asymptotically admissible. For still other sets of mean values, it is shown that no asymptotically subminimax procedure exists which is also asymptotically admissible. (This research was sponsored by the Office of Naval Research.)

17. **Note on the Problem of Combining Independent Tests of Significance.** ALLAN BIRNBAUM, Columbia University.

The "problem of combining independent tests of significance" is stated and illustrated. A distinction is made between two kinds of problems of combining tests. A definition of admissibility of a method of combining tests is proposed. A simple characterization of a minimal essentially complete class of methods of combining tests for a large class of problems is given (in terms of the proposed definition of admissibility). For a simple common testing problem, the properties of several methods of combining tests which have been recommended in the literature are examined and compared with properties of the optimal test for this problem. It is shown that some of these methods correspond to inadmissible test procedures for the original problem; other methods correspond to admissible procedures, and on the basis of their properties one may base a selection of a method of combination of tests, for use in problems which seem similar to that under consideration

18. **On Judging all Contrasts in the Analysis of Variance. (Preliminary Report.)** HENRY SCHEFFÉ, Columbia University.

Under the usual (general linear hypothesis) assumptions, if the conventional $F$-test of the hypothesis $\mathbf{H}$: $\theta_1 = \theta_2 = \cdots = \theta_k$ at the $\alpha$ level of significance rejects $\mathbf{H}$, what further conclusions are valid about the contrasts among the $\theta_i$ (beyond the conclusion that the contrasts are not all zero)? Suppose the $F$-test has $k - 1$ and $\nu$ degrees of freedom. For any $c_1, \cdots, c_k$ with $\Sigma_1^k c_i\theta_i = 0$ write $\psi_c$ for the contrast $\Sigma_1^k c_i\theta_i$, and write $\hat{\psi}_c$ and $\hat{\sigma}^2(\hat{\psi}_c)$ for the usual estimates of $\psi_c$ and the variance of $\hat{\psi}_c$. Then for the totality of contrasts $\psi_c$ the probability is $1 - \alpha$ that they all satisfy (*) $|\psi_c - \hat{\psi}_c| \leqq S\,\hat{\sigma}(\hat{\psi}_c)$, where $S^2$ is $(k - 1)$ times the upper $\alpha$ point of the $F$-distribution with $k - 1$ and $\nu$ degrees of freedom. Say that the estimated contrast $\hat{\psi}_c$ is "significantly different from zero" if the interval (*) does not cover the point $\psi_c = 0$. Then the $F$-test rejects $\mathbf{H}$ if and only if some $\hat{\psi}_c$ are significantly different from zero, and if it does, one can say just which $\hat{\psi}_c$. More generally, the inequalities (*) can be employed for all the contrasts with the obvious frequency interpretation about the proportion of experiments in which all statements are correct. Relations are considered to an earlier method of Tukey using the Studentized range tables ("Allowances for various types of error rates," invited address before IMS meeting at Blacksburg, Va., March 19, 1952; see also *Proc. Fifth Annual Convention A.S.Q.C.*, 1951, pp. 189–197). (Work sponsored by the Office of Naval Research.)

19. **Further Moving Average Methods in the Estimation of LD-50.** B. M. BENNETT, University of Washington.

Further extensions of the moving average methods for estimating LD-50 proposed by W. R. Thompson and the author (*Jour. Hyg.*, 1952) have been developed in order to include minimum variance and minimum variance unbiassed estimates from any assumed distribution of threshold. Numerical estimates of the efficiency of such procedures are included.

20. **Nonparametric Theory: Confidence Regions and Tests for Location and Scale Parameters. (Preliminary Report.)** D. A. S. FRASER, University of Toronto.

All unbiased estimates of zero are characterized for the class of probability density functions having the $p$ percentile at the origin. As a consequence, results are obtained for location parameters (percentiles), for scale parameters (interpercentile ranges), and for combination location and scale parameters. All randomized $\beta$ level confidence regions

for a location parameter can be described in a simple functional form. All confidence bounds for a location parameter are almost everywhere equal to the order statistics chosen with fixed probabilities; all distribution-free upper tolerance limits are of the same form, the Robbins condition, $\Pi(f(x_1, \cdots, x_n) - x_i) = 0$ almost everywhere, being necessary but not sufficient (unless $f$ is continuous). Most powerful one-sided and most powerful (shortest) unbiased confidence regions are obtained for the location parameters. Correspondingly, most powerful and most powerful unbiased tests are obtained for the location parameter. For the scale parameter similar or exact confidence regions do not exist. Analogously, similar tests do not exist for the scale parameter. For the scale and location parameter, confidence regions can be obtained; most powerful intervals exist for specific purposes. If the location parameter is known UMV unbiased estimates do not exist for any other parameter.

21. **Random Functions Satisfying Certain Linear Relations. I.** S. G. GHURYE, University of North Carolina.

Let $X(t)$ be a random function of the real parameter $t$, and let there exist $k$ continuous functions $\alpha_1(h), \cdots, \alpha_k(h)$, of $h > 0$, such that for any $t$ and any $h > 0$, the sequence $\{Y(t, h; n), n = 0, \pm 1, \cdots\}$, defined below, satisfies certain assumptions about independence. Here $Y(t, h; n) = X(t + (n + k)h) + \alpha_1(h)X(t + (n + k - 1)h) + \cdots + \alpha_k(h)X(t + nh)$. In other words, it is postulated that any sequence of observations made at equidistant $t$-points on the $X$-process satisfies a hypothesis of the type commonly expressed in terms of linear stochastic difference equations. Certain consequences of this assumption are studied in this paper. For instance, in all the cases considered, it is shown that the possible forms of the $\alpha_j(h)$ are restricted by the property that the roots of the following equation in $x$ can be written in the form $e^{\lambda_i h}$, the $\lambda_i$ being independent of $h$: $x^k - \alpha_1(h)x^{k-1} - \cdots - \alpha_k(h) = 0$. It is also shown that, under certain conditions, the $Y$-sequence is $(k - 1)$-dependent for all $h$, but can be independent for all $h$ only if the $X$-process is deterministic.

22. **Computational Experience in Solving Linear Programs.** A. J. HOFFMAN, National Bureau of Standards.

This is a report of the experience gained in solving linear programs on the SEAC (National Bureau of Standards Eastern Automatic Computer) by the simplex method, relaxation, and fictitious play.

23. **Duality of Multi-valued Means.** THEODORE S. MOTZKIN, University of California, Los Angeles.

In problems of statistical measurement and estimation, especially of angles and similar geometric quantities, and in their application including data preparation for linear programming, the question comes up how to define the mean for periodic variables. For these, as the example of equidistant points on a circumference shows, a unique and rotationally invariant mean cannot be defined. A multivalued and invariant mean is obtained as follows. For a set $S$ on a smaller arc the arithmetic mean is ordinarily found by choosing a zero point on the complementary arc $A_1$. It is reasonable to invest this mean with a weight proportional to the length of $A_1$, and the means obtained from zeros on the arcs $A_2, \cdots, A_n$ with corresponding weights. One may try repeating the procedure with respect to the set $S'$ of weighted means in view of possible convergence: however it turns out that $S'' = S$. More generally for every circular distribution $S$ the analogous definition of a mean by a Stieltjes-Lebesque integral leads to a dual or complementary distribution $S'$, and $S'' = S$. Similar facts hold on the torus (two periodic variables). On a sphere, even local associativity cannot be upheld; obtaining the mean of a finite set from the means of its finite subsets by iteration and passing to the limit, or as a minimum for the sum of the squares of spherical distances, are two transcendental (essentially infinitely-valued) definitions.

**24. A Note on a Dynamic Leontief Model with Substitution.** GEORGE B. DANTZIG, The Rand Corporation.

In an important class of discrete dynamic Leontief models with substitution, an optimal program for the entire time span can be obtained by determining locally the optimal basis for the first time period alone, followed by determining the optimal basis for the second time period given the first, etc. It follows that the relative weights in the objective function of performing the various activities in any time period *do not affect the choice* of activities in earlier time periods although, of course, they may affect the levels of activities in the early time periods. Thus, optimal bases for a sequence of periods are found in the order: Period 1, Period 2, Period 3, $\cdots$, after which activity levels are found by ordinary back-solution for linear equations in the order Period $n$, Period $n-1$, Period $n-2$, $\cdots$.

**25. Confidence Intervals on the Slopes of Regression Lines when Both Variables are Subject to Error.** C. O. JUNGE, JR., University of Washington.

In the problem of linear regression when both variables are subject to error, Reiersol has shown that if the errors and the "steady" parts are normally distributed then the slope of the desired regression line is not identifiable. If the errors in the two variables are independent, however, a conservative confidence interval on the slope of the regression line can be determined even though the slope is not identifiable. If the "steady" parts are a set of fixed but unknown values, independent estimates of the rank of the steady parts are sufficient to determine a meaningful confidence interval. In case an instrumental variable as defined by Geary is available, a confidence interval can be determined.

**26. Statistical Estimation from Time and Population Samples.** JULIUS A. JAHN, Washington State College.

Given a population of individuals, $I{:}1, \cdots, N_I$, and a variable $V_i$ defined over this population, a random sampling procedure is to be designed to provide unbiased estimates of the mean value of $V_i$ for the population with a minimum variance of the estimated values for certain fixed sample sizes and other cost conditions. When the variable is dependent upon time, the use of estimates from samples enumerated during one period of time to apply to later periods of time can lead to large and increasing bias due to changes with time. The solution proposed in this paper is to sub-divide the total period of time into a series of conveniently short intervals, $T{:}1, \cdots, N_T$, and to define as the universe to be sampled, the set of points $(it)$ formed by the combinations of the identification numbers for time intervals $(t)$ and those for the individuals in the population $(i)$. Correspondingly, the variable is defined over these points $(V_{it})$ in such a way that the sum over the time intervals for any one individual is equal to $V_i$. A sample design incorporating the use of weekly time intervals as "strata" within which were selected a random sample of points in each strata and a constant sampling rate for each strata was applied to derive estimates of the mean value of the total expenditures of students at the University of Washington during the Winter Quarter of 1951.

**27. The Analysis of Samples from Mixed Populations. (Preliminary Report.)** D. G. CHAPMAN, University of Washington.

Samples from natural populations will frequently include a mixture of several races or age groups. The analysis of such a sample into its component parts is useful for the determination of growth rates, mortalities, or the characteristics of the mixed populations. Several cases arise according to whether or not a simple age identifying factor is available, or whether or not supplementary information is available from other samples. Neyman's extended $\chi^2$ formulae can be applied to these problems to analyze the mixed samples into their component parts, and to estimate characteristics of the mixed population. Auxiliary problems that arise in utilizing the supplementary information are considered.

28. **Some Applications of the Log-Normal Distribution to the Study of Survival Times in Chronic Fatal Diseases, Particularly Leukemia.** HAROLD TIVEY, University of Oregon Medical School.

Illustrations of some of the types of disease processes in which a log-normal distribution will closely approximate the distribution of survival times will be presented. A summary of survival data on over 2600 cases of leukemia taken from the literature will be briefly presented, with particular emphasis on the statistical problems arising from this investigation. The application of the maximum likelihood method of Boag will be discussed.

29. **Field Trial Problems.** G. A. BAKER, University of California, Davis.

Theoretical and empirical examinations of the correspondence of certain field trials to conventional mathematical models have been made. The correspondence may be quite good but on the other hand serious defections do exist. It seems apparent that a careful study of field plot design and methods of analysis are necessary in order to obtain more accurate measures of plant performance on definite plots. The classical mathematical models have led to marked advances in the theory and practice of field trials and no attempt is being made here to belittle their importance. The point that is being made is that when finer distinctions between varieties and treatments are attempted even minor deviations from mathematical theory are important. Every effort must be made to improve investigative tools. This can be done in two ways. One is to develop more realistic models. This often fails because of lack of detailed knowledge of uniformity trials, inability to predict the behavior of biological material in complex situations, and because of mathematical difficulties. The other general method of improving field trials is to arrange the plots, select the plants, modify overall conditions, or transform the data so that classical models will clearly indicate differences.

30. **Population Forecasts, State of Washington, 1950–1960: Methodological Summary.** CALVIN F. SCHMID, VINCENT A. MILLER AND WARREN E. KALBACH, University of Washington.

The cohort-survival method was used to prepare the population forecasts for the State of Washington, April 1, 1955 and April 1, 1960, based on detailed data covering: (1) an age-sex breakdown of the population on which projections are to be based, (2) age-sex-specific mortality trends, (3) age-specific fertility trends for females in child-bearing ages, (4) age-sex-specific migration trends. Estimates are made of the age-sex distribution of the population as of the forecast dates.

31. **On the Addition of Chi-Squares. (Preliminary Report.)** JACK BORSTING, University of Oregon.

An example is used to compare the power of three tests of a simple hypothesis specifying the cell frequencies in discrete population. The tests are as follows; the chi-square test based on a single sample, a chi-square test based on the addition of five independent chi-squares from separate samples, and a test which rejects the hypothesis if any one of the several chi-squares is significantly large. For the example considered the power of the single sample test was considerably greater than the power of the other two tests. For example, for one alternative the power of the single sample test is 0.41 and the power of the test using the sum of chi-squares is 0.20. Curves are drawn comparing the powers for a set of alternatives.

32. **Decisions Based on Incomplete Information.** NORMAN DALKEY, The Rand Corporation. •

It is shown that the usual requirements of consistency, continuity, and convexity in choices made under incomplete information have the consequence that the decision maker must be guided by a set of weights, analogous to probabilities, for the possible relevant states of nature.

33. **Method of Synthesizing an Autonomous Shipbuilding Industry for an Inter-industry Input-Output Model.** CAPTAIN J. E. HAMILTON, George Washington University.

The paper presents a data-gathering and reduction program and methods of computation to construct a synthetic shipbuilding industry to provide a vector of inputs for a portion of the national bill of goods. This bill of goods is treated as an autonomous industry for application to a Leontief Input-Output matrix. The purpose of this is to test the Inter-Industry technique as to its suitability for testing the feasibility of military programs. The work consisted of assembling data sufficient to construct a suitable number of "prototype" bills of material for complete ships; of establishing a program of shipbuilding with sufficient breadth to embrace any industrial impact during the four year period under construction; of classifying the units in the program by similarities; of computing multipliers for material and manpower to pantograph the type to one of the prototypes; of developing a formula for computing manpower requirements by time periods; of programming computing machinery to compute material requirements by time periods; of calculating a vector of overhead inputs and a method of computation to include impacts of overhead; and of aggregating material, labor, and overhead into a single time-phased vector of total inputs from each of 192 industries.

34. **A Mathematical Model of the Shipbuilding Industry.** R. F. WILLIAMS, George Washington University.

This paper is a model of task of data gathering involved in an analysis of the industry-by-industry requirements in a "ship-building program." Such a program, in general, includes a variety of tasks other than actual building of ships, e.g. ship repair, ship conversion, etc. The set of data required for a program, **P**, is a matrix of numbers, $\{X_{ji}\}$, such that each $X_{ji}$ is a *measure* of the material from the $j$th one of the $n$ industrial segments serving the shipbuilding industry, required by the SBI during a certain (the $i$th) time segment for the program, **P**. First, as the raw data are necessarily found on a ship-task to ship-task basis, these must be amalgamated to arrive at the matrix $\{X_{ji}\}$. Thus an inquiry was made into the various possible means of amalgamation and a formalization of these techniques is presented. Secondly, the only practical such measure has been found to be the unit known as the "fixed dollar." Therefore, as the *cost* of a ship-task includes cost of material and cost of labor, a special inquiry is made into their respective records of data: bills of goods and wage scales, and a formal analysis of these is presented.

35. **Distribution-free Tests of Fit for Continuous Distribution Functions.** Z. W. BIRNBAUM, University of Washington.

This expository address deals with distribution-free techniques for deciding whether a sample $x_1, x_2, \cdots, x_n$ of a one-dimensional random variable was obtained from a population with a completely specified continuous cumulative distribution function $H(x)$. A number of such techniques are described and discussed. It is observed that all of them are based on statistics of the form $\phi[H(x_1), \cdots, H(x_n)]$ where $\phi(u_1, \cdots, u_n)$ is a sym-

metric function. The concepts of a distribution-free statistic and of a strongly distribution-free statistic are introduced, and the theorem is stated that, under fairly general assumptions a strongly distribution-free statistic must be of the form $\phi[H(x_1), \cdots, H(x_n)]$. A program is outlined for introducing various metrics in the space of distribution functions and for investigating the power of distribution-free tests with respect to these metrics.

**36. Acceptance Inspection by Variables when the Measurements Are Subject to Errors with Known Variance.** EDWARD A. FAY AND JOHN E. WALSH, U. S. Naval Ordnance Test Station, China Lake.

The situation considered is that where measurement of the characteristic of interest is not exact but subject to noticeable error. The measurement error is assumed to be independent of the true value of the characteristic measured and to be unbiased with known variance. The variance of the true values of the characteristic measured is assumed to have a known lower limit which is greater than zero. The probability distributions involved are assumed to be normal while the true values and measurement errors each form a random sample. For suitable specified acceptable and unacceptable fractions defective, this paper presents one-sided acceptance inspection criteria which have the property that the producer and consumer risks have specified upper bounds. Two-sided criteria can be obtained by appropriate utilization of the one-sided results.

**37. A Two-Sided Criterion for Acceptance Inspection by Variables.** EDWIN L. CROW, U. S. Naval Ordnance Test Station, China Lake.

It is desired to test whether a submitted lot has an acceptably small proportion defective, where an item is considered defective if a measurable characteristic is either less than a prescribed number or greater than another prescribed number. Two approximate solutions making use of measurements on each item of a sample of fixed size have been given by Wallis and Arnold (*Selected Techniques of Statistical Analysis*, Columbia University Press, 1947, pp. 52–57, 63-64). One of these simply combines two one-sided criteria and yields a triangular acceptance region in the $(\bar{x}, s)$ plane in the usual notation. The other corresponds to an acceptance region within the above triangle. The OC curves (probability of acceptance vs. lot proportion defective) of these criteria for particular values of the parameters are calculated by graphical integration. They depend on the lot mean. By interpolation between bounding ones of these OC curves it is possible to find an acceptance region intermediate to the two mentioned above for which the OC curve varies inappreciably with the lot mean. This two-sided criterion is generalized to the case in which the measurements are subject to error.

**38. Acceptance Inspection by Variables in the Presence of Measurement Error.** EDWARD A. FAY, U. S. Naval Ordnance Test Station, China Lake.

The lot acceptance criterion $\bar{x} + ks \leqq U$ for sampling inspection by variables (where $U$ is the upper limit to acceptable quality) is generalized to allow for measurement error. If $m$ measurements are made on each item of a sample of size $N$, the generalized criterion if of the form $\bar{x} + ks_b \leqq U$, where $s_b$ is the standard deviation of the $N$ item means. Under the assumption that the ratio $r$ of standard deviation of measurement error to standard deviation of product variability is known, a straightforward generalization of the derivation of Wallis (*Selected Techniques of Statistical Analysis*, Columbia University Press, 1947, pp. 59–61) leads to the appropriate value of $k$ for the case that $m$, $N$, and one point on the OC curve are preassigned, and to the appropriate values of $k$ and $N$ in the case that $m$ and two points on the OC curve are preassigned. It is shown that measurement error increases the sample size required to achieve the same OC curve in the ratio $1 + 2r^2/[m(k^2 + 2)]$, where $k$ is the value appropriate to the case of no measurement error.

Consequences of departure from the assumption of known $r$ are investigated by comparing OC curves.

39. **On Continuous Inspection Plans. (Preliminary Report.)** GERALD J. LIEBERMAN, Stanford University.

This paper presents two extensions of the first continuous sampling plan devised by H. F. Dodge (*Annals of Math. Stat.*, Vol. 14 (1943), pp. 264–279). Dodge's scheme provided for 100% inspection until $i$ consecutive non-defective units were observed, at which time partial inspection was instituted, namely, one out of $1/f$ units was inspected. As soon as a defect was observed, 100% inspection was reinstated. It is proven in this paper that whether the process is in a state of statistical control or not, an AOQL is assured. In fact, it is shown that always AOQL $\leqq (1/f - 1)/(i + 1/f)$. The second problem considered was that of obtaining a continuous sampling plan for the individual characteristics constituting an item. This paper presents a continuous inspection scheme similar to Dodge's for $r$ such characteristics, where $r$ is any finite integer. It guarantees an overall AOQL provided the process is in a state of statistical control. The procedure is as follows: units are divided into segments of $i$ items as they are manufactured. The $k^{\text{th}}$ characteristic ($k = 1, 2, \cdots, r$) is inspected 100% until the preceding full segment is found free of defects in characteristic $k$, at which time partial inspection is instituted, namely, one out of $1/f$ units is inspected for characteristic $k$. Full inspection is reinstated for any charactistic as soon as a defect in that characteristic is observed.

40. **An Acceptance Procedure with Two-Sided Specification Limits.** GEORGE J. RESNIKOFF, Stanford University.

A two-sided acceptance test in sampling inspection by variables is devised. The test is based on an estimate of the proportion defective which is optimum in the sense of minimum variance among unbiased estimates. In the one-sided case this estimate leads to the well-known test $\bar{x} + ks \leqq U$, where $U$ is a single upper limit defining quality. Graphically the test is a region in the $(\bar{x}, s)$ plane. The probability of a sample point $(\bar{x}, s)$ falling in the region is a function of the lot mean as well as of the proportion defective, but the resulting band of OC curves is so narrow, for the cases investigated, as to enable it to be considered a single curve insofar as industrial applications are concerned. Numerical investigation indicates that this property is itself virtually independent of the sample size. Since the acceptance region is inherently associated with the one-sided test, $\bar{x} \pm ks$, the OC curve of this one-sided test may be used as an approximation to the OC band of the two-sided test.

## NEWS AND NOTICES

*Readers are invited to submit to the secretary of the Institute news items of interest*

### Personal Items

Mr. Richard Berger, formerly with General Aniline & Film Corporation, is a Sales Analyst with General Dyestuff Corporation, New York City and is an Instructor in Statistics at Rutgers University.

Professor D. A. Darling, on leave from the Department of Mathematics, University of Michigan, is Visiting Assistant Professor of Mathematical Statistics at Columbia University for the academic year 1952–53.

Miss Marjorie Easterbrook, who recently received her M.A. degree from the University of Michigan, has accepted a position in the Research Department of the Holley Carburetor Company, Van Dyke, Michigan.

Dr. N. L. Johnson of University College, London will be engaged in teaching and research in the Department of Mathematical Statistics of the University of North Carolina for the year 1952–1953.

Mr. Edward C. Molina was awarded the Honorary Doctor of Science Degree by the Newark College of Engineering.

Dr. Ingram Olkin, formerly research assistant at the University of North Carolina, accepted an assistant professorship in Mathematics at Michigan State College.

Dr. Howard Raiffa, formerly at the University of Michigan, has been appointed Assistant Professor of Mathematical Statistics at Columbia University.

Professor Herbert Robbins of the University of North Carolina has been awarded a Fellowship by the John Simon Guggenheim Memorial Foundation and will spend the year 1952–53 as a member of the Institute for Advanced Study.

B. J. Winer has accepted a post-doctoral Ford Foundation Fellowship at the Institute of Mathematical Statistics of the University of North Carolina. The fellowship was set up for the purpose of having the incumbent study the possible applications of recent advances in mathematical statistics to experimental design in psychology. The Department of Psychology at North Carolina is co-sponsor of the fellowship.

Dr. T. O. Yntema has recently been appointed Vice President of Finance at Ford Motor Company, Dearborn, Michigan.

During the academic year of 1952–53, the following members of the Institute will be in residence at Cornell University: Robert E. Bechhofer, I. Blumen, W. T. Federer, C. R. Henderson, Mark Kac, Jack Kiefer, P. J. McCarthy, Garnet E. McCreary, J. E. Morton, T. A. Ryan, A. Schultz, Milton Sobel, R. G. D. Steel, and Lionel Weiss. Professor Bechhofer will be on leave from Columbia University to serve as Director of the Statistical Laboratory in the newly formed Statistics Center at Cornell. Dr. Milton Sobel will work with him on an Army Air Force Research Project. R. G. D. Steel, formerly of the Mathematics Department, University of Wisconsin has accepted a position as Associate Professor in the Biometrics Unit of the New York State College of Agriculture. Professor Lionel Weiss will be on leave from the University of Virginia to teach in the Department of Mathematics during the absence of Professor J. Wolfowitz, who will be on leave during the coming year.

---

Announcement has been made of competition for Fulbright awards for university lecturing and post-doctoral level research of Americans in Austria, Belgium and Luxembourg, Denmark, Egypt, France, Greece, Iraq, Italy, Japan,

Netherlands, Norway, Pakistan, Turkey, Union of South Africa, United Kingdom and Colonial Dependencies. Application must be made to Conference Board of Associated Research Councils, Committee on International Exchange of Persons, 2101 Constitution Avenue, Washington 25, D. C., no later than October 15, 1952.

## Preliminary Actuarial Examinations

### Prize Awards

The winners of the prize awards offered by the Society of Actuaries to the nine undergraduates ranking highest on the score of Part 2 of the 1952 Preliminary Actuarial Examinations are as follows:

*First Prize of $200*

| | |
|---|---|
| Freimer, Marshall | Harvard University |
| Lew, John S. | Yale University |

*Additional Prizes of $100*

| | |
|---|---|
| Bauman, Norman | Harvard University |
| Guinane, J. E. | Brown University |
| Horowitz, Michael | Cornell University |
| Kerr-Lawson, Angus | University of Toronto |
| Krause, Ralph M. | Harvard University |
| Muckenhoupt, Benjamin | Harvard University |
| Potthoff, Richard | Swarthmore College |

Ordinarily only one $200 prize is offered, but this year two candidates tied for first place. Each candidate was awarded a $200 prize.

The Society of Actuaries has authorized a similar set of prizes for the 1953 examinations based on the Part 2 scores.

The Preliminary Actuarial Examinations consist of the following three examinations:

Part 1. *Language Aptitude Examination.*
(Reading comprehension, meaning of words and word relationships, antonyms, and verbal reasoning.)

Part 2. *General Mathematics Examination.*
(Algebra, trigonometry, coordinate geometry, differential and integral calculus.)

Part 3. *Special Mathematics Examination.*
(Finite differences, probability and statistics.)

The 1953 Preliminary Actuarial Examinations will be prepared by the Educational Testing Service and will be administered by the Society of Actuaries at centers throughout the United States and Canada in May 1953. The closing date for applications is March 15, 1953.

Detailed information concerning the Examinations can be obtained from:

The Society of Actuaries
208 South LaSalle Street
Chicago 4, Illinois.

## New Members

*The following persons have been elected to membership in the Institute*

(March 3 to May 30, 1952)

**Acker, Robert D.,** B.S. (St. Peters College, Jersey City), Research Assistant in Underwriting Studies, Actuarial Division, Metropolitan Life Insurance Company, 1 Madison Avenue, New York 10, New York, also Graduate Student at New York University and Columbia University.

**Aggarwal, Om P.,** M.A. (Delhi Univ.), Research Assistant, Department of Statistics, Stanford University, Stanford, California.

**Anderson, Sigurd L.,** B.S.E. (Cornell Univ.), Graduate Student, Department of Experimental Statistics. North Carolina State College, Raleigh, North Carolina.

**Bell, Charles B., Jr.,** M.S. (Notre Dame), Research Laboratory Analyst, Operation Design Evaluation Group, Engineering Department, Santa Monica Division, Douglas Aircraft Company, *1726 W. 24th Street, Los Angeles, California.*

**Brewer, Robert A.,** B.S. (Univ. of Michigan), *12857 Longacre, Detroit 27, Michigan.*

**Canu, Jean-Francois,** M.Sc. (Univ. of Chicago), Mathematician, B.O.Q., Naval Proving Ground, Dahlgren, Virginia.

**Clemans, Kermit G.,** M.A. (Univ. of Minnesota), Instructor in Mathematics, Department of Mathematics, University of Oregon, Eugene, Oregon.

**Cox, Paul C.,** A.M. (Univ. of New Mexico), Analytic Statistician, White Sands Proving Ground, Las Cruces, New Mexico, *P. O. Box 43, Las Cruces, New Mexico.*

**David, Herbert A.,** B.Sc. (Sydney, Australia), Graduate Student, University College, London, *26 Curzon Road, London, N. 10, England.*

**Diamond, Earl L.,** A.B. (Univ. of Miami), Graduate Student, Department of Mathematical Statistics, University of North Carolina, *Room 227, Connor Dormitory, Chapel Hill, North Carolina.*

**Exselsen, Robert M.,** Ph.B. (Univ. of Chicago), Graduate Student, Department of Mathematical Statistics, Columbia University, *468 Riverside Drive, New York 27, New York.*

**Fosler, Edward A.,** B.A. (Wayne University, Michigan), Quality Control Engineer, *2206 John B., Hazel Park, Michigan.*

**Gardner, Leland A., Jr.,** B.A. (Columbia Univ.), Graduate Student, Department of Mathematical Statistics, Columbia University, New York 27, New York, *3081 Edwin Avenue, Apt. 1-E, Fort Lee, New Jersey.*

**Gordon, Mary A.,** Ph.D. (Columbia Univ.), Research Psychologist, United States Air Force, Lackland Air Force Base, *128 Main Plaza, San Antonio, Texas.*

**Graybill, Franklin,** Ph.D. (Iowa State College), Assistant Professor, Mathematics Department, Oklahoma A. and M. College, Stillwater, Oklahoma.

**Greenberger, Charles B.,** Student, Wayne University, *18971 Santa Rosa, Detroit, Michigan.*

**Guttman, Irwin,** B.Sc. (McGill Univ., Canada), Graduate Student, Toronto University, *16 Sussex Avenue, Toronto, Ontario, Canada.*

**Hadden, Elizabeth S.,** A.B. (Univ. of Mich.), Student, University of Michigan, *413 Betsy Barbour, Ann Arbor, Michigan.*

**Hall, William J.,** A.M. (Univ. of Mich.), Research Assistant and Graduate Student, Department of Mathematical Statistics, Institute of Statistics, University of North Carolina, Chapel Hill, North Carolina.

**Hamblen, John W.,** A.B. (Indiana Univ.), Graduate Assistant and Graduate Student, Mathematics Department, Purdue University, West Lafayette, Indiana.

**Harris, Bernard,** B.B.A. (City College of New York), Mathematical Statistician, Census Bureau, also Graduate Student, George Washington University, Washington 19, D. C.

**Hunter, John S.,** M.S. (North Carolina State), Graduate Assistant and Graduate Student, Institute of Statistics, *P.O. Box 5457, College Station, Raleigh, North Carolina.*

**Hurst, Frederick V., Jr.,** B.S. (Wagner College), Graduate Student, Department of Mathematical Statistics, Columbia University, New York 27, *82 Van Brunt Street, Staten Island 12, New York.*

**Kallianpur, Gopinath B.,** Ph.D. (Univ. of North Carolina), Lecturer and Research Associate, Statistical Laboratory, University of California, Berkeley 4, California.

**Laurent, Andre. G.,** Certificate of Probability and Statistics (Paris), Administrator, Institute of Statistics and Economics, *141 Rue du Ranelagh, Paris XVI, France.*

**Lopez, Alvaro,** Degree in Engineering (Escuela National de Minas, Colombia, South America), Statistician, Asociacion Nacional de Industriales, Medellin, Colombia, South America.

**Loural, Edgard de Almeida,** Civil Engineer (Univ. of Brasil), Aircraft Maintenance Engineer, Panair of Brasil, Aeroporto Santos Dumont, Rio de Janeiro, *Rua Marques de Abrantes, 189 Apt. 605, Rio de Janeiro, D.F., Brasil.*

**MacKay, John H.,** M.A. (State Univ. of Iowa), Assistant Professor of Accounting (on leave of absence), College of Commerce, Tulane University, New Orleans, (at present) Graduate Student, Department of Mathematical Statistics, University of North Carolina, *106 Church Street, Chapel Hill, North Carolina.*

**Marascuilo, Leonard A.,** B.A. (Univ. of Minnesota), Student and Teaching Assistant, Biostatistics Department, University of Minnesota, *602 North Hazel Street, St. Paul 6, Minnesota.*

**Marshall, Ross D.,** M.A. (Univ. of Mich.), Graduate Student, Department of Mathematics, University of Michigan, *915 Green St., Ann Arbor, Michigan.*

**McAdams, H. T.,** B.Ed. (Illinois State Normal), Research Chemist, Aluminum Research Laboratories, 3300 Missouri Avenue, E. St. Louis, *564 Logan Street, Bethalto, Illinois.*

**Mesner, Dale M.,** M.S. (Northwestern Univ.), Graduate Assistant, Department of Mathematics, Michigan State College, East Lansing, Michigan.

**Morse, Norman,** B.A. (Cornell Univ.), Teaching Assistant and Graduate Student, Statistics Department, New York State School of Industrial and Labor Relations, Cornell University, Ithaca, New York.

**Ohman, Carl R.,** B.A. (Knox College), Graduate Student and Assistant in Research, Department of Mathematics, Princeton University, Princeton, New Jersey.

**Ortega, Jose A.,** Agricultural Engineer (Havana Univ.), Economic Adviser of Ministry of Commerce of Cuba and Economic Adviser of Cuban's Works Confederation, *Carmen No. 157, entre Tenerife y Monte, Havana, Cuba.*

**Pachares, James,** B.S. (Univ. of Akron), Graduate Student, Department of Mathematical Statistics, University of North Carolina, Chapel Hill, North Carolina.

**Pelto, Chester R.,** M.S. (Penna. State College), Chief, Petrographic Section, M.R.D. Laboratory, Corps of Engineers, U. S. Army, *2008 Spencer Street, Omaha 10, Nebraska.*

**Rao, V. K. R. V.,** Ph.D. (Cambridge Univ.), Dean, Faculty of Social Sciences; Professor and Director, Delhi School of Economics, University of Delhi, Delhi 9, India.

**Resnikoff, George J.,** M.S. (Stanford Univ.), *1054 Myrtle Street, Palo Alto, California.*

**Roeloffs, Robert,** B.S. (Columbia School of Engineering), Associate Physicist, U. S. Steel Co., Research Laboratory; Graduate Student, Department of Engineering Statistics, Columbia University, New York, *38-15 Fifty-second Street, Long Island City 4, New York.*

**Rohloff, Albert C.,** M.S. (Purdue Univ.), Teaching Assistant and Graduate Student, Department of Mathematics, Purdue University, W. Lafayette, Indiana.

**Severo, Norman C.,** B.A. (Univ. of Buffalo), Graduate Student, Department of Mathematical Statistics, *Porter Hall 119E, Carnegie Institute of Technology, Pittsburgh, Pennsylvania.*

**Simmons, Leo D.,** M.S. (Louisiana State Univ.), Mathematician, Research Department, United Gas Corporation, Shreveport 92, Louisiana.

**Smith, F. Beckley, Jr.,** M.S. (Carnegie Institute of Tech.), Teaching Assistant, Carnegie Institute of Technology, *264 Atlanta Drive, Pittsburgh 28, Pennsylvania.*

**Sprott, David,** B.A. (Univ. of Toronto), *167 Glen Road, Toronto, Canada.*

**Steinberg, Leonard M.,** B.A. (Univ. of Toronto), Graduate Student, Department of Mathematics, University of Toronto, *54 St. Patrick St., Toronto 2B, Ontario, Canada.*

**Stubbs, Harold L.,** M.A. (Columbia Univ.), Research Associate, Electronic Research Project, Northwestern University, also Graduate Student, Department of Mathematics, Boston University, *355 So. Main Street, Sharon, Massachusetts.*

**Terzuoli, Andrew J.,** M.S. (New York Univ.), Instructor, Department of Mathematics, Polytechnic Institute of Brooklyn, 85 Livingston Street, Brooklyn 2, New York.

**Tung, Lloyd H.,** B.A. (Lingnan Univ.), Graduate Student, Department of Mathematics, University of Michigan, *6621 Gomberg, South Quadrangle, Ann Arbor, Michigan.*

**Walter, John R.,** B.A. (Univ. of Toronto), Graduate Student, Department of Mathematics, Division of Actuarial Science, University of Toronto, *8 Norma Cres., Toronto, Ontario, Canada.*

**Weiss, Irving,** M.A. (Columbia Univ.), Graduate Student, Department of Statistics, Stanford University, *Building 415, Apt. 2, Stanford Village, Stanford, California.*

**Whittlesey, John R. B.,** M.S. (Calif. Inst. of Tech.), Graduate Student, Department of Mathematical Statistics, University of North Carolina, *210 Ransom Street, Chapel Hill, North Carolina.*

**Wiggins, Alvin D.,** A.B. (Univ. of Calif.), Graduate Student, Department of Mathematics, University of California, Berkeley, *4100 Day Ave., Apt. 2C, Richmond 6, California.*

**Zelen, Marvin,** M.A. (Univ. of N. Carolina), Mathematical Statistician, Stevens Institute of Technology, *67-60 150th St., Kew Gardens Hills 67, New York.*

## REPORT OF THE EUGENE MEETING OF THE INSTITUTE

The fifty-second meeting of the Institute of Mathematical Statistics was held at the University of Oregon, June 19–21, 1952, in conjunction with the annual meeting of the Biometric Society, WNAR, the Eugene meetings of the Econometric Society, the Western Regional meeting of the American Mathematical Society and the Pacific Northwest meeting of the Mathematical Association of America. One hundred seventy-three persons registered, including the following members of the Institute:

C. B. Allendoerfer, Hjalmar Anderson, Jr., G. A. Baker, B. M. Bennett, C. A. Bennett, Mrs. C. A. Bennett, Z. W. Birnbaum, Jack Borsting, A. H. Bowker, Bernice Brown, D. G. Chapman, K. G. Clemans, E. L. Cox, E. L. Crow, G. B. Dantzig, R. C. Davis, W. J. Dixon, Robert Dorfman, Mary Elveback, E. A. Fay, R. S. Gardner, M. A. Girshick, C. H. Gordon, Margaret Gurney, K. D. C. Haley, W. C. Healy, Jr., H. M. Hughes, Mark Kac, R. B. Leipnik, J. C. R. Li, G. J. Lieberman, Eugene Lukacs, F. J. Massey, Jr., Paul Meyer, R. H. Moore, Sigeiti Moriguti, W. L. Nicholson, Donald Owen, Edward Paulson, Bayard Rankin, G. J. Resnikoff, Herman Rubin, Whitney Scobert, Paul Simpson, W. B. Simpson, Rosedith Sitgreaves, Andrew Sobczyk, D. Teichroew, F. H. Tingey, R. H. Titman, Elizabeth Vaughan, L. H. Wegner, Irving Weiss.

The opening session on Thursday morning was on *Stochastic Models.* Professor M. A. Girshick, Stanford University was chairman. The speakers and their subjects were:

1. *The Auditory Cortex—A Probability Model.* Archie R. Tunturi, University of Oregon Medical School.
2. *Testing Message Diffusion: The Utilization of Mathematical Models.* Stuart C. Dodd, Richard J. Hill, and Susan Huffaker, University of Washington.

This was followed by two invited addresses. Professor Frank J. Massey, Jr., University of Oregon presided. The papers were:

3. *On the Method of Collective Marks.* Herman Rubin, Stanford University.
4. *On Asymptotic Properties of Estimates.* Lucien LeCam, University of California, Berkeley.

On Friday afternoon Professor Albert H. Bowker, Stanford University presided at a session at which the following contributed papers were presented:

1. *The Principle of Invariance Applied to a Problem of Classification.* Rosedith Sitgreaves, Stanford University.
2. *Optimum selection procedures.* Edward Paulson, University of Washington.
3. *A Theorem on Convex Cones with Applications to Linear Inequalities.* Jerry W. Gaddum, Institute for Numerical Analysis.
4. *On the Prediction of Nonstationary Stochastic Processes.* R. C. Davis, Bureau of Ordnance.
5. *Power Efficiency Function for Normal Alternatives for Several Nonparametric Tests.* W. J. Dixon, University of Oregon.
6. *Bounds for Second Moments of the Sample Range.* Sigeiti Moriguti, University of North Carolina and University of Tokyo.
7. *Estimates, Tests, and Tolerance Intervals for Lump Data from a Normal Distribution.* Harry M. Hughes, University of California, Berkeley.
8. *A Computing Formula for the Power of the Two-sided t-test for Even Degrees of Freedom.* Wesley L. Nicholson, University of Oregon.
9. *On the Analysis of Data Matched in Pairs.* (By title.) Frank J. Massey, Jr., University of Oregon.
10. *The Power Function of the Haldane-Smith Test.* (By title.) B. M. Bennett, University of Washington.
11. *Spectral Analysis of Stationary Time Series. Preliminary Report.* (By title.) Ulf Grenander and Murray Rosenblatt, University of Chicago.
12. *Asymptotically Subminimax and Asymptotically Admissible Statistical Decision Procedures.* (By title.) Jack Laderman, Columbia University.
13. *Note on the Problem of Combining Independent Tests of Significance.* (By title.) Allan Birnbaum, Columbia University.
14. *On Judging all Contrasts in the Analysis of Variance. Preliminary report.* (By title.) Henry Scheffé, Columbia University.
15. *Further Moving Average Methods in the Estimation of LD-50.* (By title.) B. M. Bennett, University of Washington.
16. *Nonparametric Theory: Confidence Regions and Tests for Location and Scale Parameters. Preliminary Report.* (By title.) D. A. S. Fraser, University of Toronto.
17. *Random Functions Satisfying Certain Linear Relations. I.* (By title.) S. G. Ghurye, University of North Carolina.

Professor Robert Dorfman, University of California presided at the opening

session Friday. This session was devoted to *Developments in Linear Programming* and the following papers were presented:

1. *Computational Experience in Solving Linear Programs.* A. J. Hoffman, National Bureau of Standards.
2. *Duality of Multi-valued Means.* Theodore S. Motzkin, University of California, Los Angeles.
3. *A Note on a Dynamic Leontief Model with Substitution.* George B. Dantzig, The Rand Corporation.

Later in the morning following an address of welcome to the University by President Harry K. Newburn, Professor Mark Kac, Cornell University, presented an address titled *On Some Discrete Non-Markovian Processes.*

Friday afternoon Professor Mary Elveback, University of California, Los Angeles, presided at a session at which the following contributed papers were presented.

1. *Confidence Intervals on the Slopes of Regression Lines when both Variables Are Subject to Error.* C. O. Junge, Jr., University of Washington.
2. *Statistical Estimation from Time and Population Samples.* Julius A. Jahn, Washington State College.
3. *The Analysis of Samples from Mixed Populations. Preliminary Report.* D. G. Chapman, University of Washington.
4. *Some Applications of the Log-Normal Distribution to the Study of Survival Times in Chronic Fatal Diseases, Particularly Leukemia.* Harold Tivey, University of Oregon Medical School.
5. *Field Trial Problems.* G. A. Baker, University of California, Davis.
6. *Population Forecasts, State of Washington, 1950–1960: Methodological Summary.* Calvin F. Schmid, Vincent A. Miller and Warren E. Kalbach, University of Washington.
7. *On the Addition of Chi-squares. Preliminary Report.* Jack Borsting, University of Oregon.

On Saturday morning Professor Paul B. Simpson, University of Oregon presided at a session on *Decision Theory* at which the following three papers were presented:

1. *Decisions Based on Incomplete Information.* Norman Dalkey, The Rand Corporation.
2. *Method of Synthesizing an Autonomous Shipbuilding Industry for an Interindustry Input-Output Model.* Captain J. E. Hamilton, George Washington University.
3. *A Mathematical Model of the Shipbuilding Industry.* R. F. Williams, George Washington University.

Saturday afternoon, Professor Z. W. Birnbaum, University of Washington presented a Special Invited Paper, *Distribution-free Tests of Fit for Continuous Distribution Functions.* Professor D. G. Chapman, University of Washington, presided.

The final session was on *Inspection Sampling.* Professor Albert Bowker, Stanford University, was chairman and the following papers were presented:

1. *Acceptance Inspection by Variables when the Measurements Are Subject to Errors with Known Variance.* Edward A. Fay and John E. Walsh, U. S. Naval Ordnance Test Station, China Lake.

2. *A Two-sided Criterion for Acceptance Inspection by Variables*. Edwin L. Crow, U. S. Naval Ordnance Test Station, China Lake.
3. *Acceptance Inspection by Variables in the Presence of Measurement Error*. Edward A. Fay U. S. Naval Ordnance Test Station, China Lake.
4. *On Continuous Inspection Plans. Preliminary Report*. Gerald J. Lieberman, Stanford University.
5. *An Acceptance Procedure with Two-sided Specification Limits*. George J. Resnikoff, Stanford University.

Social events included a banquet sponsored by the mathematical organizations, followed by an open house at the Faculty Club Friday evening.

W. J. DIXON
*Assistant Secretary*

## PUBLICATIONS RECEIVED

ABRUZZI, ADAM, *Work Measurement*, Columbia University Press, New York, 1952, xviii + 290 pp., $6.00.

MELTZER, HANS, *Wahrheit und Wahrscheinlichkeit in der Statistik*, Wirtschaftschoehschule Mannheim, 1952, iv + 35 pp.

*Quaderni dell' Istituto di Scienze Economiche e Statistiche*, xiv, Università degli Studi di Milano, Milan, 1951, iv + 52 pp., L. 400.

*Recenseamento Geral do Brasil* (1° de Setembro de 1940) *Censo Demográfico* and *Censos Econômicos*, Serviço Gráfico de Instituto Brasileiro de Geográfia e Estatística, Rio de Janeiro, 1950, 12 volumes in addition to those listed in March, 1952.

*Statistical Method in Industrial Production*, (Papers given at a Conference held in Sheffield in 1950), Royal Statistical Society, London, 1951, v + 89 pp., 7 shillings and sixpence.

WAUGH, ALBERT E., *Elements of Statistical Method*, McGraw-Hill, New York, 1952, xv + 531 pp., $5.50

# ESTADISTICA

*Journal of the Inter American Statistical Institute*

**Vol. X, No. 36** **September 1952**

***Contents***

Estadística No. 36 is an experimental "subject issue," devoted almost in its entirety to:

STATISTICAL TEACHING AND TRAINING
METHODS AND MATERIALS

This number was developed primarily to provide a collection of useful viewpoints and important reference materials which would be of particular value to statistical teaching and training activities in the Western Hemisphere—whether conducted in universities, in government statistical operating agencies, or elsewhere. The contents fall roughly into five categories: (1) Problems and needs in statistical education; (2) how the needs can be met; (3) examples of curricula and operational methods; (4) reports and other documentary material; and (5) bibliography.

The material is published under three separate covers—the main *Journal*, consisting of approximately 170 pages, in which articles in their original text appear; *Supplement No. 1*, consisting of approximately 120 pages of translations of articles or reports, and of other documentary material; and *Supplement No. 2*, containing a 50-page Bibliography of Statistical Textbooks and Other Teaching Material, prepared under the auspices of the Committee on Statistical Education of IASI.

Among contributors to issue No. 36 are the following: FRANCISCO DE ABRISQUETA, CARLOS E. DIEULEFAIT, HALBERT L. DUNN, FORREST E. LINDER, M. PEREZ GUERRERO, STUART A. RICE, LUIS THORIN CASAS, SAMUEL S. WILKS, and others.

**Published quarterly; annual subscription price $3.00 (U.S.); single copies $1.00 (U.S.)**

**Inter American Statistical Institute, % Pan American Union, Washington 6, D.C., U. S. A.**

---

## JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

***June 1952***

**Vol. 47 No. 258**

**1108 16th Street, N. W., Washington 6, D. C.**

An Analysis of Some Failure Data ..... D. J. DAVIS
Classification and Analysis of Partially Balanced Incomplete Block Designs with Two Associate Classes ..... R. C. BOSE AND T. SHIMANOTO
Replacing Variables in Correlation Problems ..... VINCENT I. WEST
Operating Characteristics for Tests of the Stability of a Normal Population ..... JOHN E. WALSH
Multiple Sampling of Attributes ..... D. S. ROBSON AND A. J. KING
Estimating the Product of Several Random Variables ..... G. D. SHELLARD
Some Principles of Processing Census and Survey Data ..... ROBERT B. VOIGHT AND MARTIN KRIESBERG
Short-cut Methods of Estimating County Population ..... ROBERT C. SCHMITT
Sampling Surveys in Central Africa ..... J. R. H. SHAUL
Factors in the Accumulation of Social Statistics ..... SOLOMON FABRICANT
Comparison of Selected Measures of Ability of Communities to Bear Tax Burdens ..... LORIN A. THOMPSON
A Critical Evaluation of Available Agricultural Statistics ..... IVAN M. LEE
Prepaid Medical Care as a Source of Morbidity Data ..... NEVA R. DEARDORFF
Some Cases in Which Yates' Correction Should not be Applied ..... EDWIN L. CROW

**ASA**

The American Statistical Association invites as members all persons interested in:

1. development of new theory and method
2. improvement of basic statistical data
3. application of statistical methods to practical problems.

# BIOMETRIKA

*A Journal for the Statistical Study of Biological Problems*

**Volume 39** **Contents** **Parts 3 and 4, December 1952**

1. Estimation of population parameters from data obtained by means of capture-recapture method. Part II. By P. H. LESLIE. 2. Tensor notation and the sampling cumulants of k-statistics. By E. L. KAPLAN. 3. Estimation in double sampling. By D. R. COX. 4. Sampling from bivariate, non-normal universes. By H. HYRENIUS. 5. The truncated Poisson distribution. By P. G. MOORE. 6. Upper 5% and 1% points of the ratio $s^2max/s^2min$. By H. DAVID. 7. Conditions under which Gram-Charlier and Edgeworth curves are positive definite and unimodal. By D. E. BARTON and K. E. DENNIS. 8. On a two-sided sequential t-test. By S. RUSHTON. 9. Properties of distribution based on certain simple transformation of the normal curve. By J. DRAPER. 10. Estimation of the mean and standard deviation of a normal population from a censored sample. By A. K. GUPTA. 11. The rank analysis of incomplete block designs. By R. A. BRADLEY and M. E. TERRY. 12. The statistical structure of ecological communities. By J. G. SKELLAM. 13. The growth, survival, wandering and variation of the long-tailed field mouse. Part III—Wandering and recapture. By H. P. and H. S. HACKER. 14. Use of scores for the analysis of association in contingency tables. By E. J. WILLIAMS. 15. Statistical significance of odd bits of information. By M. S. BARTLETT. 16. Tests of fit in time series. By P. WHITTLE. 17. The fitting of grouped truncated and grouped censored normal distributions. By P. M. GRUNDY. 18. MISCELLANEA: Samples with the same number in each stratum. By W. L. STEVENS. Comparison of analysis of variance power functions in the parametric and random models. By N. L. JOHNSON. Approximation to the probability integral of the distribution of range. By N. L. JOHNSON. Discrimination in time series analysis. By A. RUDRA. Statistical control of counting experiments. By H. O. LANCASTER. Exact grouping corrections to moments and cumulants. By M. KUPPERMAN.

**The subscription price, *payable in advance*, is 45s. inland, 54s. export (per volume including postage). Cheques should be drawn to Biometrika and sent to "The Secretary, Biometrika Office, Department of Statistics, University College, London, W.C. 1." All foreign cheques must be in sterling and drawn on a bank having a London agency.**

# ECONOMETRICA

***Journal of the Econometric Society***

***Contents of Vol. 20, July, 1952, include:***

JAN TINBERGEN
Four Alternative Policies to Restore Balance of Payments Equilibrium
T. M. BROWN.....Habit Persistence and Lags in Consumer Behavior
A. D. ROY.....Safety First and the Holding of Assets
WALTER ISARD.....A General Location Principle of an Optimum Space-Economy
RENÉ ROY.....Les Elasticites
de la Demande relative aux biens de Consommation et aux groups de biens
A. DVORETZKY, J. KIEFER, AND J. WOLFOWITZ
The Inventory Problem: II. Case of Unknown Distributions of Demand
Reports of the Boston and Tokyo Meetings
Review Articles, Book Reviews, and Notes

---

Published Quarterly **Subscription rates available on request**

The Econometric Society is an international society for the advancement of economic theory in its relation to statistics and mathematics.

Subscriptions to *Econometrica* and inquiries about the work of the Society and the procedure in applying for membership should be addressed to William B. Simpson, Secretary, The Econometric Society, The University of Chicago, Chicago 37, Illinois, U. S. A.

# MATHEMATICAL REVIEWS

*A journal containing reviews of the mathematical literature of the world, with full subject and author indices*

Publication of this journal is sponsored by the American Mathematical Society, Mathematical Association of America, Institute of Mathematical Statistics, London Mathematical Society, Edinburgh Mathematical Society, Union Matematica Argentina, and others.

Subscriptions accepted to cover the calendar year only.
Issues appear monthly except July. $20.00 per year.

***Send subscription order or request for sample copy to***

**AMERICAN MATHEMATICAL SOCIETY**
**80 Waterman Street, Providence 6, Rhode Island**



## JOURNAL OF THE
# ROYAL STATISTICAL SOCIETY

***Series B (Methodological)***

***Contents of Volume 13, No. 2, 1951***


DAVID G. KENDALL.....Some Problems in the Theory of Queues
H. E. DANIELS.....The Theory of Position Finding. (With Discussion)
W. R. BUCKLAND.....A Review of the Literature of Systematic Sampling
V. P. GODAMBE.....On Two-stage Sampling
B. V. SUKHATME...On Certain Probability Distributions Arising from Points on a Line
P. G. GUEST..The Estimation of Standard Error from Successive Finite Differences
E. H. SIMPSON.....The Interpretation of Interaction in Contingency Tables
H. O. PATTERSON.....Complex Contingency Tables Treated by the Partition of $\chi^2$
S. ROSENBAUM.....The Variance of Least-Square Estimates under Linear Restraint
H. D. PATTERSON.....Change-over Trials
P. M. GRUNDY
A General Technique for the Analysis of Experiments with Incorrectly Treated Plots
D. J. FINNEY..Subjective Judgment in Statistical Analysis: An Experimental Study
G. H. JOWETT The Expression of the Complementary
Outputs of Two Products in Terms of a Common Unit of Production Effort
J. DURBIN AND A. STUART.....Inversions and Rank Correlation Coefficients
H. E. DANIELS.....Note on Durbin and Stuart's Formula for $E(r_s)$


**The Royal Statistical Society, 4, Portugal Street, London, W.C.2.**

# SKANDINAVISK AKTUARIETIDSKRIFT

**1951 – Parts 3 – 4**

***Contents***


G. ARFWEDSON
A Probability Distribution Connected with Stirling's Second Class Numbers
T. DALENIUS AND M. GURNEY ..... The Problem of Optimum Stratification. II.
HERMAN WOLD ..... Demand Functions and the Integrability Condition
G. HARBITZ ..... Post-war Mortality among Industrial Insured Lives in Norway
S. VAJDA ..... Analytical Studies in Stop-Loss Insurance
KNUT MEDIN
The Sickness Experience of the Valkyrian Insurance Company, 1929–1948
PER OTTESTAD
On the Test of the Hypothesis that the Probability of an Event Is Contained within Given Limits


**Annual subscription: 10 Swedish Crowns (Approx. $2.00).**

**Inquiries and orders may be addressed to the Editor.**

**SKÄRVIKSVAGEN 7, DJURSHOLM (SWEDEN)**

# SANKHYĀ

***The Indian Journal of Statistics***

***Edited by P. C. Mahalanobis***

**Vol. 11, Part 2, 1951**


The Estimation of Parameters in Certain Stochastic Processes .... HENRY B. MANN
Statistical Inference Applied to Classificatory Problems .... C. RADHAKRISHNA RAO
Multivariate Binomial and Poisson Distributions ..... A. S. KRISHNAMOORTHY
On Errors of Estimates in Various Types of Double Sampling Procedure .. K. C. SEAL
Estimation of Parameters from Incomplete Data with Application to Design of Sample Surveys ..... ABRAHAM MATTHAI
Confluent Hypergeometric Function ..... PRAN NATH
A Study of Recent Trend in Infantile Mortality Rates in Calcutta by Longitudinal Survey ..... K. N. MITRA, B. BHATTACHARYA, K. DEY, C. S. DAWN, M. OBADIAH, AND A. K. GAYEN
On the Non-Existence of Certain Difference Sets for Incomplete Group Designs S. S. SHRIKHANDE
On the Non-Existence of Affine Resolvable Balanced Incomplete Block Designs S. S. SHRIKHANDE
A Note on the Power of the Best Critical Region for Increasing Sample Size D. BASU
Some Further Results on Errors in Double Sampling Technique .... CHAMELI BOSE
A Note on Price-Wage Variations in Cottage and Factory Economy .. G. C. MANDAL


**Annual subscription: 30 rupees**
**Inquiries and orders may be addressed to the**
**Editor, Sankhyā, Presidency College, Calcutta, India.**